

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
16 June 2005 (16.06.2005)

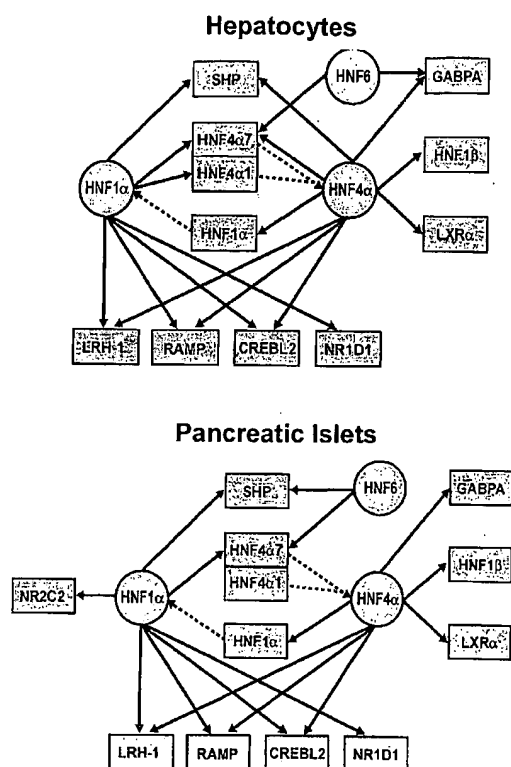
PCT

(10) International Publication Number
WO 2005/054461 A2

- (51) International Patent Classification⁷: C12N 15/00
- (21) International Application Number: PCT/US2004/039805
- (22) International Filing Date: 23 November 2004 (23.11.2004)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
60/525,318 26 November 2003 (26.11.2003) US
60/542,520 6 February 2004 (06.02.2004) US
60/544,835 13 February 2004 (13.02.2004) US
60/547,933 26 February 2004 (26.02.2004) US
- (71) Applicant (for all designated States except US): **WHITE-HEAD INSTITUTE FOR BIOMEDICAL RESEARCH** [US/US]; Nine Cambridge Center, Cambridge, MA 02142-1479 (US).
- (72) Inventors; and
- (73) Inventors/Applicants (for US only): **ODOM, Duncan, T.** [US/US]; 48 JFK St. Apt. 20, Cambridge, MA 02138 (US). **YOUNG, Richard, A.** [US/US]; 216 Highland Street, Weston, MA 02193 (US).
- (74) Agents: **GRANAHAN, Patricia** et al.; Ropes & Gray LLP, One International Place, Boston, MA 02110-2624 (US).
- (81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

[Continued on next page]

(54) Title: TRANSCRIPTIONAL REGULATORS AND METHODS THEREOF



(57) Abstract: The invention relates to transcriptional regulators and related methods thereof. The invention further relates to the identification of genes regulated by transcriptional regulators, to the treatment of diseases associated with abnormal function of a transcriptional regulator and to the modulation of gene expression, including genes expressed in hepatocytes or pancreatic cells, through the modulation of transcriptional regulator activity.



(84) **Designated States** (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

Transcriptional Regulators and Methods Thereof

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims the benefit of the filing date of U.S. Application No. 60/525318, filed November 26, 2003, entitled "CONTROL OF PANCREAS AND LIVER GENE EXPRESSION BY HNF TRANSCRIPTION FACTORS", U.S. Application No. 60/542520, filed February 6, 2004, entitled "CONTROL OF PANCREAS AND LIVER GENE EXPRESSION BY HNF TRANSCRIPTION FACTORS", U.S. Application No. 60/544835, filed February 13, 2004, entitled "CONTROL OF PANCREAS AND LIVER GENE EXPRESSION BY HNF TRANSCRIPTION FACTORS", and U.S. Application No. 60/547933, filed February 26, 2004, entitled "TRANSCRIPTIONAL REGULATORS AND METHODS THEREOF". The entire teachings of the referenced applications are incorporated by reference herein.

FUNDING

The invention described herein was supported, in whole or in part, by the U.S. Department of Energy Program for Computational Molecular Biology. The United States government has certain rights in the invention.

BACKGROUND OF THE INVENTION

Gene expression is controlled by transcriptional regulatory proteins, which bind specific DNA sequences and recruit cofactors and the transcription apparatus to promoters (1-3). The expression of transcriptional regulators themselves is also regulated by transcriptional regulators, and a single gene may be regulated by multiple transcription factors. As a result of these regulatory networks, or pathways, misregulation of a single transcriptional regulator in a cell can result in the aberrant expression of multiple genes in the network in which the transcriptional regulator is active, leading to disease in the organism.

Current methods of identifying the genes controlled by a transcriptional regulator typically include a comparison of the mRNA levels of candidate target in

cells which express the transcriptional regulator and control cells which either do not express it. Often, this involves overexpressing a recombinant transcriptional regulator in a given cell type and using, as a control cell, one which overexpresses a control recombinant protein or no recombinant protein at all. However, given to the artificial
5 nature of using cell lines and overexpressing transgenes, the results obtained from such approaches may not reflect the *in vivo* regulation by native transcriptional regulators in an organism.

Genome-wide analysis methods have been used recently to determine how
10 tagged transcriptional regulators encoded in *Saccharomyces cerevisiae* are associated with the genome in living yeast cells and to model the transcriptional regulatory circuitry of these cells (4). These methods have also been used in human tissue culture cells to identify target genes for several transcriptional regulators (5-7).

15 However, the need remains to develop genome-scale analysis methods to determine how transcriptional regulators control the global gene expression programs that characterize specific tissues, and in particular, freshly isolated, primary tissues, in which the transcriptional regulators are likely to maintain their *in vivo* specificities. Furthermore, there is a need to identify the regulatory networks or pathways in which a
20 given transcriptional activator acts, in part, to allow for the identification of therapeutic targets for diseases caused by aberrant function of a transcriptional regulator.

SUMMARY OF THE INVENTION

In one aspect, the invention provides a method of identifying the genes
25 regulated by a transcriptional regulator. One aspect of the invention provides a method of determining which genes from a subset of genes are regulated by a transcriptional regulator in a cell, the method comprising (a) selectively isolating chromatin from a cell which expresses the transcriptional regulator to generate isolated chromatin; (b) selectively isolating chromatin fragments from the isolated chromatin to generate
30 bound chromatin fragments, wherein the bound chromatin fragments are bound by the transcriptional regulator; (c) amplifying both the bound chromatin fragments to generate amplified chromatin fragments and the isolated chromatin to generate

amplified control chromatin; (d) hybridizing the amplified control chromatin and the amplified chromatin fragments to a DNA microarray, wherein the DNA microarray comprises (1) at least 10,000 experimental spots, each experimental spot comprising an experimental DNA, each experimental DNA comprising a promoter region from a gene
5 in the subset; and (2) at least 100 control spots, each control spot comprising a control DNA, each control DNA comprising a non-promoter region; and (e) determining and comparing a hybridization signal at each of the spots on the microarray between those generated by (1) the amplified control chromatin; and (2) the amplified chromatin fragments; wherein a gene in the subset is said to be regulated by the transcriptional
10 regulator in the cell if a spot comprising a promoter region of said gene displays a higher level of hybridization by the amplified chromatin fragments than by the amplified control chromatin.

In another aspect, the invention provides methods of identifying regulatory
15 networks, or pathways, in a cell. The invention provides a method of identifying a transcriptional regulatory network in a cell, the method comprising determining if a transcriptional regulator regulates additional transcriptional regulators in the cell using the method of any of the methods described herein, wherein a transcriptional regulatory network is identified if at least one additional transcriptional regulator is
20 regulated by the transcriptional regulator.

The invention also provides a method of identifying a transcriptional regulatory network in a cell, the method comprising determining if a transcriptional regulator regulates (i) its own promoter; or (ii) a promoter from a plurality of transcriptional
25 regulators; using any of the methods described herein, wherein the experimental DNA comprises (a) a promoter from the transcriptional regulator; and (b) promoters from the plurality of transcriptional regulators;
wherein a transcriptional regulatory network is identified if the transcriptional regulator regulates itself or if it regulates at least one of the plurality of transcriptional regulators.

30

The invention further provides a method of identifying transcriptional regulatory networks in a cell, the method comprising (a) determining, by repeating a

method of identifying the targets of transcriptional regulator for each of a plurality of transcriptional regulators, the genes in a subset which are regulated by each of the plurality of transcriptional regulators, wherein the experimental DNA comprises promoter regions for each of the plurality of transcriptional regulators; (b) determining
5 if any one of the plurality of transcriptional regulators are regulated by at least one of the plurality of transcriptional regulators; wherein a transcriptional regulatory network is identified if any one of the plurality of transcriptional regulators is regulated by at least one of the plurality of transcriptional regulators.

10 The invention also provides a DNA microarray for determining promoter occupancy in a human cell, the microarray comprising (1) at least 10,000 experimental spots, each experimental spot comprising an experimental DNA, each experimental DNA comprising a promoter region from a human gene in the subset; and (2) at least 100 control spots, each control spot comprising a control DNA, each control DNA
15 comprising a non-promoter region; wherein at least 75% of the promoter regions comprise from at least 700bp upstream to at least 200 bp downstream of the transcriptional start site.

Another aspect of the invention provides a method of estimating if a
20 transcriptional regulator is a global transcriptional regulator, the method comprising (a) selectively isolating chromatin from a tissue; (b) identifying promoter regions from the chromatin which are bound by a candidate global transcriptional regulator; (c) identifying promoter regions from the chromatin which are bound by a member of the basal transcriptional machinery; and (d) comparing the promoter regions identified
25 in steps (b) and (c) to determine the ratio between (i) the number of promoter regions bound by both the candidate global transcriptional regulator and the member of the basal transcriptional machinery; and (ii) the number of promoter regions bound by the member of the basal transcriptional machinery, wherein a transcriptional regulator is a global transcriptional regulator when the ratio is greater than 0.2.

30

The invention further provides methods of identifying targets for therapeutics. In one aspect, the invention provides a method of identifying at least one target gene for

the development of a therapeutic to treat or prevent a disorder in a subject, wherein at least one form of the disorder is caused by an altered activity in a transcriptional regulator or in a suspected transcriptional regulator, the method comprising (a) identifying the genes regulated by the transcriptional regulator in a cell; (b) 5 determining if the transcriptional regulator is a broad-acting transcriptional regulator or a narrow-acting transcriptional regulator, wherein if the transcriptional regulator is a broad acting transcriptional regulator then the transcriptional regulator is a target gene for the development of a therapeutic, and wherein if the transcriptional regulator is a narrow acting transcriptional regulator then (i) determining if at least one gene 10 regulated by the transcriptional regulator is likely causative in the disorder, wherein a gene that is likely causative in the disorder is a target gene for the development of a therapeutic; and (ii) reiterating steps (a) and (b) for at least one gene that is regulated by the transcriptional regulator in the cell and that either (1) encodes a transcriptional regulator or (2) is suspected to encode a transcriptional regulator, with the modification 15 that the transcriptional regulator of steps (a) and (b) is said gene, thereby identifying at least one target gene for the development of a therapeutic to treat or prevent a disorder in the subject.

The invention also provides methods of treating or preventing disease. In one 20 aspect, the invention provides a method of treating or preventing type II diabetes in a subject, comprising administering to the subject a therapeutically effective amount of an agent that increases the global transcriptional activity of HNF4alpha.

In another aspect, the invention provides a method of treating or preventing a 25 disorder associated with low transcriptional activity of HNF4alpha in a subject, comprising administering to the subject a therapeutically effective amount of an agent that increases the global transcriptional activity of HNF4alpha. A related aspect provides a method of treating or preventing a disorder associated with high transcriptional activity of HNF4alpha in a subject, comprising administering to the 30 subject a therapeutically effective amount of an agent that decreases the global transcriptional activity of HNF4alpha.

The invention also provides a method of increasing the global transcriptional activity in a liver or a pancreatic cell comprising contacting the cell with an agent which increases the global transcriptional activity of HNF4alpha. A related aspect provides a method of decreasing the global transcriptional activity in a liver or a
5 pancreatic cell comprising contacting the cell with an agent which decreases the global transcriptional activity of HNF4alpha.

One aspect of the invention provides methods of regulating the expression level of genes. On aspect provides a method of regulating the expression level of any one of
10 the genes in Figure 13 in a hepatocyte, the method comprising contacting the cell with an agent which regulates the transcriptional activity of HNF1alpha. A related aspect provides a method of regulating the expression level of any one of the genes in Figure 14 in a pancreatic cell, the method comprising contacting the cell with an agent which regulates the transcriptional activity of HNF1alpha.

15 Another aspect of the invention provides a method of regulating the expression level of any one of the genes in Figure 16 in a hepatocyte, the method comprising contacting the cell with an agent which regulates the transcriptional activity of HNF6. A related aspect provides a method of regulating the expression level of any one of the
20 genes in Figure 17 in a pancreatic cell, the method comprising contacting the cell with an agent which regulates the transcriptional activity of HNF6.

Yet another aspect of the invention provides a method of regulating the expression level of any one of the genes in Figure 18 in a hepatocyte, the method
25 comprising contacting the cell with an agent which regulates the transcriptional activity of HNF4alpha. A related aspect provides a method of regulating the expression level of any one of the genes in Figure 19 in a pancreatic cell, the method comprising contacting the cell with an agent which regulated the transcriptional activity of HNF4alpha.

30

The invention also provides methods for identifying transcriptionally active genes that are regulated by a transcriptional regulator in a cell. In one aspect, the

invention provides a method of identifying transcriptionally active genes that are regulated by a transcriptional regulator in a cell, the method comprising (a) selectively isolating chromatin from a tissue; (b) identifying promoter regions from the chromatin that are bound by the transcriptional regulator; (c) identifying promoter regions from the chromatin that are bound by a member of the basal transcriptional machinery; and (d) comparing the promoter regions identified in steps (b) and (c) to determine overlapping genes, wherein the overlapping genes are transcriptionally active genes regulated by the transcriptional regulator.

10 BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A-1C show genome-scale location analysis of HNF regulators in human tissues. (A) Hepatocytes and pancreatic islets were obtained from tissue distribution programs. These cells were treated with formaldehyde to covalently link transcription factors to DNA sites of interaction. Cells were harvested, and chromatin in cell lysates was sheared by sonication. The regulator-DNA complexes were enriched by chromatin immunoprecipitation with specific antibodies, the crosslinks were reversed, and enriched DNA fragments and control genomic DNA fragments were amplified using ligation-mediated PCR. The amplified DNA preparations, labeled with distinct fluorophores, were mixed and hybridized onto a promoter array. (B) Venn diagram showing the overlap of HNF1 α , HNF6, and HNF4 α bound promoters in hepatocytes (top) and pancreatic islets (bottom). (C) The collection of genes occupied by RNA polymerase II in hepatocytes is displayed as a circle, with the genes bound by HNF1 α , HNF6, and HNF4 α outlined collectively as a fraction of the chart. The relative contributions of HNF1 α , HNF6, and HNF4 α are shown as framing arcs.

25

Figures 2A-2B show transcriptional regulatory networks and motifs. (A) HNF1 α , HNF6, and HNF4 α are at the center of tissue-specific transcriptional regulatory networks. In these examples selected for illustration, regulatory proteins and their gene targets are represented as circles and boxes, respectively. Solid arrows indicate protein-DNA interactions, and genes encoding regulators are linked to their protein products by dashed lines. The HNF4a7 promoter, also known as the P2 promoter (24, 25), was recently implicated as a major human diabetes susceptibility locus (see text). (B)

30

Examples of regulatory network motifs in hepatocytes. For instance, in the multi-component loop, HNF1 α protein binds to the promoter of the HNF4 α gene, and the HNF4 α protein binds to the promoter of the HNF1 α gene. These network motifs were uncovered by searching binding data with various algorithms; for details on the algorithms used and a full list of motifs found, see (20).

Figure 3 shows one embodiment of a strategy for the identification of at least one target gene of a master regulator for the development of a therapeutic to treat or prevent a disorder.

Figure 4 shows a Venn diagram showing the overlap of two single, independent ChIP experiments using hepatocytes with anti-HNF4 α antibodies sc-6556 and sc-8987.

Figure 5 shows a Western blot of HNF4 α in HepG2 cells using 50 μ g of cell lysate protein with Ab sc-6556. The lower running band is approximately 50 kDa, which is the canonical molecular weight for HNF4 α , and the higher running band is the appropriate location for HNF4 α dimer. A very similar gel showing HNF4 α antibody specificity for sc-6556 is available at the Santa Cruz website (www.scbt.com).

Figures 6A-6D show scatterplots of attempted chromatin immunoprecipitations performed with the anti-HNF4 α antibody sc-6556 using Jurkat (T-lymphocyte derived, 6A), BJ-T (foreskin fibroblast derived, 6B), and U937 (histocyte derived, 6C) cells. To demonstrate the noise inherent in the array analysis, applicants show a scatterplot of a sample of input DNA, split, labeled with the two fluorophores, and hybridized to an array (6D). Identical control experiments performed using the anti-HNF1 α antibody sc-6547 afforded essentially identical results.

Figure 7 shows a scatterplot of a chromatin immunoprecipitation performed with pre-immune commercial rabbit serum using hepatocytes (left). Goat pre-immune serum and two rabbit sera from different individuals gave a similar scatterplot. For comparison, applicants show the scatterplot for an equivalent ChIP with the anti-HNF4 α antibody sc-6556 using hepatocytes (right).

Figure 8 shows a Venn diagram showing the overlap of the sets of promoters bound by HNF4 α and RNA Pol II in hepatocytes and pancreatic islets.

- 5 **Figure 9** shows a composite gel of gene-specific chromatin immunoprecipitation reactions using anti-HNF4 α antibody sc-6556 with crosslinked human hepatocytes.

Figure 10 shows composite gel of gene-specific chromatin immunoprecipitation reactions using anti-HNF1 α antibody sc-6547 with crosslinked human hepatocytes.

10

Figure 11 shows a partial list of proximal promoters occupied by of HNF1a in human hepatocytes and pancreatic islets. These genes were assigned to functional categories using the program ProtoGo; genes not in this automated GO ontology database were assigned using Locuslink information. Four genes are shown for each tissue/category combination; for some combinations, fewer than 4 promoters qualified as targets.

15

Hypothetical and functionally uncharacterized genes are not shown. A complete list of targets is available in Figures 13 and 14.

Figure 12 shows Occupancy of BJ-T and tissue-specific promoter sets by HNF factors.

20

(*) Indicates that comparisons between BJ-T and primary tissues used only a subset of Hu13K array promoters, as RNA Pol II was profiled in BJ-T cells using a smaller, prototype array. The denominator in the above fractions represents the number of targets the HNF factor of interest occupied in the set of RNA Pol II occupied promoters that are either BJ-T specific or primary tissue specific.

25

Figure 13 shows HNF1 α bound promoters in hepatocytes

Figure 14 shows HNF1 α bound promoters in pancreatic islets.

- 30 **Figures 15A-15D** show genes previously suggested to be regulated by HNF1a and HNF4a. 'Direct' binding is in vivo ChIP and in vivo footprinting, 'in vitro' binding is primarily gel mobility retardation assays and in vitro footprinting, and 'indirect' is

primarily transient transfections. 'Sequence-based' uses a number of different criteria to qualify binding. Note that some duplicate reports are omitted, as are a handful of recent large-scale screens, (e.g. Tronche 1997, Shih 2001, etc.).

5 **Figure 16** shows HNF6 bound promoters in hepatocytes.

Figure 17 shows HNF6 bound promoters in pancreatic islets.

Figure 18A-18C show HNF4 α bound promoters in hepatocytes.

10

Figures 19A-19C show HNF4 α bound promoters in pancreatic islets.

Figures 20A-20B show the feed forward regulatory motifs in hepatocytes . The regulatory modules here were derived as described in exemplification. Feed forwards only involving HNF1 α and HNF4 α are also multi-input motifs, as they bind each other's promoters in a multicomponent loop.

15

Figures 21A-21B show multi-input motifs in hepatocytes. The regulatory modules here were derived as described in the exemplification. MIMs for the HNF6/HNF4 α and HNF1 α /HNF4 α are listed in Figure 20 as feedforward motifs.

20

Figures 22A-22B show the feed forward regulatory motifs in pancreatic islets . The regulatory modules here were derived as described in Supporting Online Material. Feed forwards only involving HNF1 α and HNF4 α are also multiinput motifs, as they bind each other's promoters in a multicomponent loop.

25

Figures 23A-23B show multi-Input motifs in pancreatic islets The regulatory modules here were derived as described in Supporting Online Material. MIMs for the HNF6/HNF4 α and HNF1 α /HNF4 α are listed in Figure 22 as feedforward.

30

Figures 24A-24B show transcriptional regulators occupied by HNF1 α and HNF4 α . Network of DNA regulators downstream of HNF1 α and HNF4 α in hepatocytes and

islets. Target genes that are among the Gene Ontology "DNA-regulators" category were compiled, and are listed according to functional subcategory.

DETAILED DESCRIPTION OF THE INVENTION

5 I. Overview

In certain aspects, the invention provides methods related to transcriptional regulators. Some aspects of the invention provide methods for the identification of genes whose transcription is regulated by a specific transcriptional regulator in a cell. Some of these methods comprise determining the promoter occupancy of the
10 transcriptional regulator using a combination of chromatin immunoprecipitation and/or DNA microarray analysis of the promoter regions that are physically associated with the transcriptional regulator in the cell. In some embodiments of the methods described herein, the DNA microarray comprises both experimental spots containing promoter DNA, and control spots containing non-promoter DNA. The methods described herein
15 may be applied to any cell type, including transplant grade primary human tissue. Furthermore, the method described herein can be used to compare the function of transcriptional regulators across cell types, or across two populations, such as healthy and disease-afflicted subjects.

20 In a related aspect, the invention provides methods of identifying regulatory networks, or pathways. Some methods comprise identifying the transcriptional regulators which are regulated by a given transcriptional regulator, and optionally, determining the genes that are regulated by those transcriptional regulators. Pathways that may be identified using the methods described herein include autoregulatory,
25 multicomponent, feed-forward, and multi-components loops, as well as regulatory chains.

The invention also provides methods of determining if a transcriptional regulator is a global transcriptional regulator. In some aspects, such methods comprise
30 determining the promoter occupancy of both a transcriptional regulator and a member of the basal transcriptional machinery. Comparison of the promoter occupancy by the transcriptional regulator and by the member of the basal transcriptional machinery

allows the identification of transcriptionally active promoters that are bound and regulated by the transcription regulator. Other methods further comprise extrapolating from the set of promoters that were examined to the total number of promoters in the genome to determine the approximate number of transcriptionally active promoters in a cell that are under the control of a specific transcriptional factor or to determine if the transcriptional regulator is a global transcriptional regulator.

Other aspects of the invention provide methods of identifying therapeutic targets to treat disease. One specific aspect of the invention relates to identifying at least one target gene for the development of a therapeutic agent to treat or prevent a disorder in a subject, preferably a disorder in which at least one form of the disorder is caused by an altered activity in a transcriptional regulator or in a gene suspected to encode a transcriptional regulator. Some of the methods provided herein to identify therapeutic targets comprise determining if a transcriptional regulator implicated in the disease is a broad-acting or a narrow-acting transcriptional regulator, such as by identifying at least a subset of the genes that it regulates in a cell, wherein broad-acting transcriptional regulators are targets for therapeutic agents. If the transcriptional regulator is narrow-acting, then the genes that it regulates may be examined further to determine if any are broad-acting transcriptional regulators (for those genes encoding transcriptional regulators) or if any of the genes are causative to the disease state *i.e.* they regulate a pathway or network that is impaired in the disease state.

The invention further provides methods for the treatment of disease. Some aspects of the invention provide methods of treating metabolic disorders, such as type II diabetes. Specific aspects of the invention provide methods of treating or preventing type II diabetes in a subject by administering to the subject a therapeutically effective amount of an agent that increases the global transcriptional activity of HNF4 α . Furthermore, the invention provides methods for modulating the expression level of genes. Such methods are based, in part, on the finding by Applicants of genes which are transcriptionally regulated by HNF1 α , HNF4 α or HNF6 in hepatocytes and pancreatic cells. In a related aspect, the invention provides methods of modulating and expression level of, and alleviating a disease state associated with the abnormal

expression of, the genes in Figures 13-19 by modulating the transcriptional activity or expression of HNF1 α , HNF4 α or HNF6. In specific embodiments, the expression of the genes is modulated in hepatocytes, pancreatic cells, or both.

5 II. Definitions

For convenience, certain terms employed in the specification, examples, and appended claims, are collected here. Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs.

10

The articles "a" and "an" are used herein to refer to one or to more than one (i.e., to at least one) of the grammatical object of the article. By way of example, "an element" means one element or more than one element.

15

The term "including" is used herein to mean, and is used interchangeably with, the phrase "including but not limited" to.

The term "or" is used herein to mean, and is used interchangeably with, the term "and/or," unless context clearly indicates otherwise.

20

The term "such as" is used herein to mean, and is used interchangeably, with the phrase "such as but not limited to".

25 A "patient" or "subject" to be treated by the method of the invention can mean either a human or non-human animal, preferably a mammal.

The terms "alpha" and " α " are used interchangeably, as are the terms "beta" and " β ".

30 The term "encoding" comprises an RNA product resulting from transcription of a DNA molecule, a protein resulting from the translation of an RNA molecule, or a protein resulting from the transcription of a DNA molecule and the subsequent

translation of the RNA product.

A "promoter" is a nucleic acid sequence that directs transcription of a nucleic acid. A promoter includes nucleic acid sequences near the start site of transcription, e.g., a TATA box, see, e.g., Butler and Kadonaga (2002) *Genes Dev.* 16:2583-2592; 5
Georgel (2002) *Biochem. Cell Biol.* 80:295-300. A promoter also optionally includes distal enhancer or repressor elements, which can be located as much as several thousand base pairs on either side from the start site of transcription. A "constitutive" promoter is a promoter that is active under most environmental and developmental 10
conditions, while an "inducible", promoter is a promoter is active or activated under, e.g., specific environmental or developmental conditions.

The term "expression" is used herein to mean the process by which a polypeptide is produced from DNA. The process involves the transcription of the gene 15
into mRNA and the translation of this mRNA into a polypeptide. Depending on the context in which used, "expression" may refer to the production of RNA, protein or both.

The term "recombinant" is used herein to mean any nucleic acid comprising 20
sequences which are not adjacent in nature. A recombinant nucleic acid may be generated *in vitro*, for example by using the methods of molecular biology, or *in vivo*, for example by insertion of a nucleic acid at a novel chromosomal location by homologous or non-homologous recombination.

25 The term "transcriptional regulator" refers to a biochemical element that acts to prevent or inhibit the transcription of a promoter-driven DNA sequence under certain environmental conditions (e.g., a repressor or nuclear inhibitory protein), or to permit or stimulate the transcription of the promoter-driven DNA sequence under certain environmental conditions (e.g., an inducer or an enhancer).

30 The term "microarray" refers to an array of distinct polynucleotides or oligonucleotides synthesized on a substrate, such as paper, nylon or other type of

membrane, filter, chip, glass slide, or any other suitable solid support.

The terms "disorders" and "diseases" are used inclusively and refer to any deviation from the normal structure or function of any part, organ or system of the body
5 (or any combination thereof). A specific disease is manifested by characteristic symptoms and signs, including biological, chemical and physical changes, and is often associated with a variety of other factors including, but not limited to, demographic, environmental, employment, genetic and medically historical factors. Certain characteristic signs, symptoms, and related factors can be quantitated through a variety
10 of methods to yield important diagnostic information.

The terms "level of expression of a gene in a cell" or "gene expression level" refer to the level of mRNA, as well as pre-mRNA nascent transcript(s), transcript processing intermediates, mature mRNA(s) and degradation products, encoded by the
15 gene in the cell.

The term "modulation" refers to upregulation (i.e., activation or stimulation), downregulation (i.e., inhibition or suppression) of a response, or the two in combination or apart. A "modulator" is a compound or molecule that modulates, and may be, e.g.,
20 an agonist, antagonist, activator, stimulator, suppressor, or inhibitor.

The term "agonist" refers to an agent that mimics or up-regulates (e.g., potentiates or supplements) the bioactivity of a protein, e.g., polypeptide X. An agonist may be a wild-type protein or derivative thereof having at least one bioactivity of the
25 wild-type protein. An agonist may also be a compound that upregulates expression of a gene or which increases at least one bioactivity of a protein. An agonist may also be a compound which increases the interaction of a polypeptide with another molecule, e.g., a target peptide or nucleic acid.

30 The term "antagonist" refers to an agent that downregulates (e.g., suppresses or inhibits) at least one bioactivity of a protein. An antagonist may be a compound which inhibits or decreases the interaction between a protein and another molecule, e.g., a

target peptide or enzyme substrate. An antagonist may also be a compound that downregulates expression of a gene or which reduces the amount of expressed protein present.

5 The term “prophylactic” or “therapeutic” treatment refers to administration to the subject of one or more of the subject compositions. If it is administered prior to clinical manifestation of the unwanted condition (e.g., disease or other unwanted state of the host animal) then the treatment is prophylactic, i.e., it protects the host against developing the unwanted condition, whereas if administered after manifestation of the
10 unwanted condition, the treatment is therapeutic (i.e., it is intended to diminish, ameliorate or maintain the existing unwanted condition or side effects therefrom).

 The term “therapeutic effect” refers to a local or systemic effect in animals, particularly mammals, and more particularly humans caused by a pharmacologically
15 active substance. The term thus means any substance intended for use in the diagnosis, cure, mitigation, treatment or prevention of disease or in the enhancement of desirable physical or mental development and conditions in an animal or human. The phrase “therapeutically-effective amount” means that amount of such a substance that produces some desired local or systemic effect at a reasonable benefit/risk ratio
20 applicable to any treatment. In certain embodiments, a therapeutically-effective amount of a compound will depend on its therapeutic index, solubility, and the like. For example, certain compounds discovered by the methods of the present invention may be administered in a sufficient amount to produce a reasonable benefit/risk ratio applicable to such treatment.

25 A probe that is “labeled” is detectable, either directly or indirectly, by spectroscopic, photochemical, biochemical, immunochemical, isotopic, or chemical means. For example, useful labels include ^{32}P , ^{33}P , ^{35}S , ^{14}C , ^3H , ^{125}I , stable isotopes, fluorescent dyes and fluorettes (Rozinov and Nolan (1998) Chem. Biol 5:713-728; Molecular Probes, Inc. (2003) Catalogue, Molecular Probes, Eugene Oreg.), electron-
30 dense reagents, enzymes and/or substrates, e.g., as used in enzyme-linked immunoassays as with those using alkaline phosphatase or horse radish peroxidase. The

label or detectable moiety is typically bound, either covalently, through a linker or chemical bound, or through ionic, van der Waals or hydrogen bonds to the molecule to be detected. "Radiolabeled" refers to a compound to which a radioisotope has been attached through covalent or non-covalent means. A "fluorophore" is a compound or moiety that absorbs radiant energy of one wavelength and emits radiant energy of a second, longer wavelength.

A "labeled nucleic acid probe or oligonucleotide" is one that is bound, either covalently, through a linker or a chemical bond, or noncovalently, through ionic, van der Waals, electrostatic, or hydrogen bonds to a label such that the presence of the probe can be detected by detecting the presence of the label bound to the probe. The probes are preferably directly labeled as with isotopes, chromophores, fluorophores, chromogens, or indirectly labeled such as with biotin to which a streptavidin complex or avidin complex can later bind.

A "nucleic acid probe" is a nucleic acid capable of binding to a target nucleic acid of complementary sequence, usually through complementary base pairing, e.g., through hydrogen bond formation. A probe may include natural, e.g., A, G, C, or T, or modified bases, e.g., 7-deazaguanosine, inosine, etc. The bases in a probe can be joined by a linkage other than a phosphodiester bond. Probes can be peptide nucleic acids in which the constituent bases are joined by peptide bonds rather than phosphodiester linkages. It will be understood by one of skill in the art that probes may bind target sequences lacking complete complementarity with the probe sequence depending upon the stringency of the hybridization conditions.

"Small molecule" is defined as a molecule with a molecular weight that is less than 10 kD, typically less than 2 kD, and preferably less than 1 KD. Small molecules include, but are not limited to, inorganic molecules, organic molecules, organic molecules containing an inorganic component, molecules comprising a radioactive atom, synthetic molecules, peptide mimetics; and antibody mimetics. As a therapeutic, a small molecule may be more permeable to cells, less susceptible to degradation, and less apt to elicit an immune response than large molecules. Small molecule toxins are

described, see, e.g., U.S. Pat. No. 6,326,482 issued to Stewart, et al.

A small molecule refers to a composition, which has a molecular weight of less than about 1000 kDa.

5 III. Identification of Transcriptional Targets and Transcriptional Networks

One aspect of the invention provides a method of determining which genes from a subset of genes are regulated by a transcriptional regulator in a cell, the method comprising (a) selectively isolating chromatin from a cell which expresses the transcriptional regulator to generate isolated chromatin; (b) selectively isolating
10 chromatin fragments from the isolated chromatin to generate bound chromatin fragments, wherein the bound chromatin fragments are bound by the transcriptional regulator; (c) amplifying both the bound chromatin fragments to generate amplified chromatin fragments and the isolated chromatin to generate amplified control chromatin; (d) hybridizing the amplified control chromatin and the amplified
15 chromatin fragments to a DNA microarray, wherein the DNA microarray comprises (1) at least 10,000 experimental spots, each experimental spot comprising an experimental DNA, each experimental DNA comprising a promoter region from a gene in the subset; and (2) at least 100 control spots, each control spot comprising a control DNA, each control DNA comprising a non-promoter region; and (e) determining and
20 comparing a hybridization signal at each of the spots on the microarray between those generated by (1) the amplified control chromatin; and (2) the amplified chromatin fragments; wherein a gene in the subset is said to be regulated by the transcriptional regulator in the cell if a spot comprising a promoter region of said gene displays a higher level of hybridization by the amplified chromatin fragments than by the
25 amplified control chromatin.

Methods of isolating chromatin, and in particular chromatin fragments that are bound by a transcriptional regulator, may be carried out by any method known to one skilled in the art, including by cross-linking the transcriptional regulator to chromatin,
30 fragmenting the chromatin, and immunoprecipitating the transcriptional regulators.

In a preferred embodiment, the chromatin fragments bound by the

transcriptional regulator are isolated using chromatin immunoprecipitation (ChIP).

Briefly, this technique involves the use of a specific antibody to immunoprecipitate chromatin complexes comprising the corresponding antigen *i.e.* the transcriptional regulator, and examination of the nucleotide sequences present in the

5 immunoprecipitate. Immunoprecipitation of a particular sequence by the antibody is indicative of interaction of the antigen with that sequence. See, for example, O'Neill et al. in *Methods in Enzymology*, Vol. 274, Academic Press, San Diego, 1999, pp. 189-197; Kuo et al. (1999) *Method* 19:425-433; and Ausubel et al., *supra*, Chapter 21.

10 In one embodiment, the chromatin immunoprecipitation technique is applied as follows. Cells which express the transcriptional regulator of interest, such as a native transcriptional regulator or a recombinant transcriptional regulator, are treated with an agent that crosslinks the transcriptional regulator to chromatin if that transcriptional regulator is stably bound to it. In one embodiment of the methods described herein, the
15 crosslinking is formaldehyde crosslinking (Solomon, M.J. and Varshavsky, A., *Proc. Natl. Sci. USA* 82:6470-6474; Orlando, V., *TIBS*, 25:99-104). UV light may also be used (Pashev et al. *Trends Biochem Sci.* 1991;16(9):323-6; Zhang L et al. *Biochem Biophys Res Commun.* 2004;322(3):705-11).

20 Subsequent to crosslinking, cellular nucleic acid is isolated, sheared such as by sonication and incubated in the presence of an antibody directed against the transcriptional regulator. Antibody-antigen complexes are precipitated, crosslinks are reversed (for example, formaldehyde-induced DNA-protein crosslinks can be reversed by heating) so that the sequence content of the immunoprecipitated DNA is tested for
25 the presence of a specific sequence, for example, promoter regions. The antibody may bind directly to an epitope on the transcriptional regulator or it may bind to a tag on the regulator, such as a myc tag when used with an anti-Myc antibody (Santa Cruz Biotechnology, sc-764).

30 In yet another embodiment, a non-antibody agent with affinity for the transcriptional regulator or for a tag used to it is used in place of the antibody. For example, if the transcriptional regulator comprises an affinity tag, such as a six-

histidine tag, complexes may be isolated by affinity chromatography to nickel-containing sepharose. Additional variations on ChIP methods within the scope of the invention may be found in Kurdistan et al. Methods. 2003 31(1):90-5; O'Neill et al. Methods. 2003, 31(1):76-82; Spencer et al., Methods. 2003;31(1):67-75; and Orlando
5 et al. Methods 11: 205-214 (1997).

In an alternate embodiment of the methods described herein for identifying genes regulated by a transcriptional regulator, amplified chromatin fragments from a control immunoprecipitation reaction are used in place of the isolated chromatin as a
10 control. For example, an antibody that does not react with the transcription factor being tested may be used in a chromatin IP procedure to isolate control chromatin, which can then be compared to the chromatin isolated using an antibody that does react with the transcriptional regulator. In preferred embodiments, the antibody that does not react with the transcription factor being tested also does not react with other transcriptional
15 regulators or DNA binding proteins.

In one embodiment, the amplified control chromatin and the amplified chromatin fragments are generated from their corresponding template DNA using ligation-mediated polymerase chain reaction (LM-PCR) (e.g., see Current Protocols in
20 Molecular Biology, Ausubel, F. M. et al., eds. 1991, and U.S. Application No. 2003/0143599, the teachings of which are incorporated herein by reference) in their entirety. In specific embodiments, LM-PCR comprises fluorescently labeling amplified DNA by including fluorescently-tagged nucleotides in the LM-PCR reaction. Additional variations for manipulating and examining chromatin using microarrays
25 have described in U.S. Patent Nos. 6,410,243, the teachings of which are incorporated herein by reference.

In one embodiment, the labelled or unlabeled probes are hybridized to DNA microarray, such as is described in U.S. Patent No. 6,410,243. Microarrays, also called
30 "biochips" or "arrays" are miniaturized devices typically with dimensions in the micrometer to millimeter range for performing chemical and biochemical reactions and are particularly suited for embodiments of the invention. Arrays may be constructed via

microelectronic and/or microfabrication using essentially any and all techniques known and available in the semiconductor industry and/or in the biochemistry industry, provided only that such techniques are amenable to and compatible with the deposition and screening of polynucleotide sequences. Microarrays are particularly desirable for
5 their virtues of high sample throughput and low cost for generating profiles and other data. Additional variations for manipulating and examining chromatin using microarrays have described in U.S. Patent Nos. 6,410,243, the teachings of which are incorporated herein by reference.

10 In one embodiment of the methods described, amplified control chromatin and the amplified chromatin fragments are hybridized to a DNA microarray that includes experimental spots that represent all or a subset (e.g., a chromosome or chromosomes) of the genome. The fluorescent intensity of each experimental spot on the microarray from the amplified chromatin fragments relative to the amplified control chromatin
15 indicates whether the protein of interest is bound to the DNA region located at that particular spot. Hence, the methods described herein allow the detection of protein-DNA interactions across an entire genome.

In some embodiments of the methods described herein, the promoter region of a
20 gene comprises from at least 700bp upstream to at least 200 bp downstream of the transcriptional start site of the gene. In some embodiments, the promoter region comprises at least about 30, 40, 50, or 60 nucleotides in length. In specific embodiments, the promoter region of a gene as found on the spots of the microarray comprises a sequence of at least 30 nucleotides whose sequence is identical to a region
25 stretching from 3 kb upstream to 1 kb downstream of the transcriptional start site of said gene. In some embodiments, the DNA microarray includes control spots of non-promoter DNA. In specific embodiment, the non-promoter region comprises an open reading frame. In preferred embodiments, the non-promoter regions comprise genomic regions which are not bound by transcriptional regulators, and preferably which are not
30 bound by the transcriptional regulator being tested. In some embodiments, not all the experimental spots or the control spots comprise experimental DNA or control DNA, respectively. Furthermore, in some specific embodiments some spots comprise control

DNA which comprises promoter DNA. One skilled in the art may determine the number of experimental or control spots for a given application.

5 In some embodiments of the methods described herein, the level of hybridization of the amplified chromatin fragments to each experimental spot is normalized by the level of hybridization of the amplified chromatin fragments to the control spots. In specific embodiments, the normalization is performed by subtracting the mean level of hybridization of the amplified chromatin fragments to the control spots from the level of hybridization of the amplified chromatin fragments at each
10 experimental spot.

Methods of analyzing data from microarrays are well-described in the art, including in DNA Microarrays: A Molecular Cloning Manual, Ed by Bowtel and Sambrook (Cold Spring Harbor Laboratory Press, 2002); Microarrays for an Integrative
15 Genomics by Kohana (MIT Press, 2002); A Biologist's Guide to Analysis of DNA Microarray Data, by Knudsen (Wiley, John & Sons, Incorporated, 2002); and DNA Microarrays: A Practical Approach, Vol. 205 by Schema (Oxford University Press, 1999); and Methods of Microarray Data Analysis II, ed by Lin et al. (Kluwer Academic Publishers, 2002), hereby incorporated by reference in their entirety.

20 In some embodiments of any of the methods described herein, the transcriptional regulator is native to the cell. By native it is meant that the transcriptional regulator naturally occurs in the cell. In other embodiments, the transcriptional regulator is a recombinant transcriptional regulator. In some
25 embodiments, the transcriptional regulator originates from a species which is different from that of the cell. In some embodiments, the transcriptional regulator is a viral transcriptional regulator. In such embodiments, a cell may be contacted with a virus and chromatin extracted from the infected cell after allowing sufficient time for the viral proteins to be expressed. In some embodiments, recombinant transcriptional
30 regulators have missense mutations, truncations, or inserted sequences or entire domains from other naturally occurring proteins. A tagged recombinant transcriptional regulator may be used in some embodiments the methods of the present invention as

the tag may facilitate the immunoprecipitation of the regulator.

In certain embodiments of the invention, transcriptional regulators comprise specific transcription factors, coactivators, corepressors or complexes thereof.

5 Transcription factors bind to specific cognate DNA elements such as promoters, enhancers and silencer elements, and are responsible for regulating gene expression. Transcription factors may be activators of transcription, repressors of transcription or both, depending on the cellular context. Transcription factors may belong to any class or type of known or identified transcription factor. Examples of known families or
10 structurally-related transcription factors include helix-loop-helix, leucine zipper, zinc finger, ring finger, and hormone receptors. Transcription factors may also be selected based upon their known association with a disease or the regulation of one or more genes. For example, transcription factors such as c-myc, Rel/Nf-kB, neuroD, c-fos, c-jun, and E2F may be targeted. Antibodies directed to any transcriptional coactivator or
15 corepressor may also be used according to the invention. Examples of specific coactivators include CBP, CTIIA, and SRA, while specific examples of corepressors include the mSin3 proteins, MITR, and LEUNIG. Furthermore, the genes regulated by proteins associated with transcriptional complexes, such as the histone acetylases (HATs) and histone deacetylases (HDACs), may also be determined using the methods
20 described herein.

In one embodiment of the methods described herein, the cell is a primary cell. Primary cells are directly isolated from an organism and have undergone minimum passaging *in vitro*, and thus maintain most of the phenotypic characteristics of cells in
25 the organism. In a specific embodiment, the primary cells are primary cells that have doubled less than 10 times *ex vivo*. In some embodiments, the cell is derived from transplant grade tissue or freshly isolated tissue. The cell type used in the assays described herein may be any cell type. The cell may be eukaryotic or prokaryotic, from a metazoan or from a single-celled organism such as yeast. In some preferred
30 embodiments the cell is a mammalian cell, such as a cell from a rodent, a primate or a human. The cell may be a wild-type cell or a cell that has been genetically modified by recombinant means or by exposure to mutagens. The cell may be a transformed cell or

an immortalized cell. In some embodiments, the cell is from an organism afflicted by a disease. In some embodiments, the cell comprises a genetic mutation that results in disease, such as in a hyperplastic condition.

5 In some embodiments, the cell is derived from transplant-grade tissue or freshly isolated tissue. In some embodiments, the cell is derived from a tissue biopsy, such as from a subject afflicted with, or suspected of being afflicted with, a disorder. In another embodiment, the cell is isolated from a bodily fluid or bodily secretion, including serum, plasma, saliva, tears, sweat, semen, amniotic fluid, vaginal secretions,
10 nasal secretions, synovial fluid, spinal fluid, phlegm, bronchoalveolar lavage fluid, blister fluid, pus, stool and intracranial fluid. The cell may be a live cell or a cell that has been preserved, such as by treatment with formalin, B5, Zenker's fixatives, Lugol's solution, Carnoy's Fixative, F13 fixative, or other preservatives, or a cell that has been preserved by freezing.

15 In some embodiments of the methods described herein, the cell has been treated with an agent, such as compound or a drug, prior to isolation of chromatin. Some preferred agents include those which bind to or regulate the expression of transcriptional regulators. In some embodiments, the genes that are regulated by a
20 given transcriptional regulator are determined both in a cell that is contacted with an agent and in a cell that is not contacted with the agent, or that is contacted with a different amount of the agent. Such methods may be used to identify compounds that alter the types of genes and/or the extent to which a transcriptional regulators controls transcription of those genes. Furthermore, such approaches may be used to screen for
25 agents which alter the activity, specificity or expression of a transcriptional regulator.

 In some embodiment of the methods described herein for identifying genes regulated by a transcriptional regulator, a higher level of hybridization by the amplified chromatin fragments than by the amplified control chromatin comprises at least a two-
30 fold higher level of hybridization. The threshold for what constitutes a higher level of hybridization, may be adjusted by one skilled in the art for the particular application. Higher levels of hybridization are expected to yield a smaller target size but with higher

certainty that a given gene above that threshold is regulated by the transcriptional regulator in that cell *in vivo*.

In other embodiments of the methods described herein for identifying genes
5 regulated by a transcriptional regulator, the transcriptional regulator is a basal
transcription factor or a component of the basal transcription machinery. In specific
embodiments, components of the basal transcription machinery comprise RNA
polymerases, including poII, poIII and poIII, TBP, NTF-1 and Sp1 and any other
component of TFIID, including, for example, the TAFs (e.g. TAF250, TAF150,
10 TAF135, TAF95, TAF80, TAF55, TAF31, TAF28, and TAF20), or any other
component of a polymerase holoenzyme.

Another aspect of the invention provides a method of identifying
transcriptionally active genes that are regulated by a transcriptional regulator in a cell.
15 The method comprises determining what genes are regulated by the transcriptional
regulator and determining which ones are transcriptionally active in the cell. In one
embodiment, a set of genes which are transcriptionally active is the set of genes whose
promoters are bound by an RNA polymerase, such as RNA polymerase II, or by a
member of the basal transcription machinery. Alternatively, genes which are
20 transcriptionally active may be identified using other techniques known in the art. For
example, mRNA from a cell which expresses the transcriptional regulator can be
collected and examined on a DNA microarray which comprises coding sequences in
order to determine which genes are being transcribed.

25 In one embodiment, the invention provides a method of identifying
transcriptionally active genes that are regulated by a transcriptional regulator in a cell,
the method comprising (a) selectively isolating chromatin from a tissue; (b) identifying
promoter regions from the chromatin that are bound by the transcriptional regulator; (c)
identifying promoter regions from the chromatin that are bound by a member of the
30 basal transcriptional machinery; and (d) comparing the promoter regions identified in
steps (b) and (c) to determine overlapping genes, wherein the overlapping genes are
transcriptionally active genes regulated by the transcriptional regulator.

In a related aspect, the invention provides methods to determine if a transcriptional regulator is a global transcription regulator. One method comprises estimating if a transcriptional regulator is a global transcriptional regulator, the method comprising (a) selectively isolating chromatin from a tissue; (b) identifying promoter regions from the chromatin which are bound by a candidate global transcriptional regulator; (c) identifying promoter regions from the chromatin which are bound by a member of the basal transcriptional machinery; and (d) comparing the promoter regions identified in steps (b) and (c) to determine the ratio between (i) the number of promoter regions bound by both the candidate global transcriptional regulator and the member of the basal transcriptional machinery; and (ii) the number of promoter regions bound by the member of the basal transcriptional machinery wherein a transcriptional regulator is a global transcriptional regulator when the ratio is greater than 0.2.

In a preferred embodiment of the methods described above, steps (b) and (c) are performed using a DNA microarray. In a specific embodiment, the DNA microarray comprises (i) at least 10,000 experimental spots, each experimental spot comprising an experimental DNA, each experimental DNA comprising a promoter region from a human gene in the subset; and (ii) at least 100 control spots, each control spot comprising a control DNA, each control DNA comprising a non-promoter region. Any type of microarray or array may be used.

In one embodiment of the methods described above, the member of the transcriptional machinery is an RNA polymerase, such as RNA polymerase II, a TATA-binding protein, or any other component of TFIID, including, for example, the TAFs (e.g. TAF250, TAF150, TAF135, TAF95, TAF80, TAF55, TAF31, TAF28, and TAF20).

Another aspect of the invention provides methods of identifying regulatory networks, or pathways, in a cell. The methods provided by the invention allow the identification of the regulatory motifs, such as those shown in Figure 2B. A regulatory pathway can include, for example, a pathway that controls a cellular function under a

specific condition. A regulatory pathway controls a cellular function by, for example, altering the activity of a system component or the activity of a biochemical, gene expression or other type of pathway. Alterations in activity include, for example, inducing a change in the expression, activity, or physical interactions of a pathway component under a specific condition. Specific examples of regulatory pathways include a pathway that activates a cellular function in response to an environmental stimulus of a biochemical system, such as the inhibition of cell differentiation in response to the presence of a cell growth signal and the activation of galactose import and catalysis in response to the presence of galactose and the absence of repressing sugars. The term "component" when used in reference to a network or pathway is intended to mean a molecular constituent of the biochemical system, network or pathway, such as, for example, a polypeptide, nucleic acid, other macromolecule or other biological molecule.

15 In one aspect, the invention provides a method of identifying a transcriptional regulatory network in a cell, the method comprising determining if a transcriptional regulator regulates additional transcriptional regulators in the cell, such as by using any of the methods described herein, wherein a transcriptional regulatory network is identified if at least one additional transcriptional regulator is regulated by the
20 transcriptional regulator.:

Another aspect of the invention provides a method of identifying a transcriptional regulatory network in a cell, the method comprising determining if a transcriptional regulator regulates (i) its own promoter; or (ii) a promoter from a plurality of transcriptional regulators; such as by using any of the methods described
25 herein, wherein the experimental DNA comprises (a) a promoter from the transcriptional regulator; and (b) promoters from the plurality of transcriptional regulators; wherein a transcriptional regulatory network is identified if the transcriptional regulator regulates itself or if it regulates at least one of the plurality of
30 transcriptional regulators.

Yet another aspect of the invention provides a method of identifying

transcriptional regulatory networks in a cell, the method comprising (a) determining, by repeating one of the methods described herein for each of a plurality of transcriptional regulators, the genes in a subset which are regulated by each of the plurality of transcriptional regulators, wherein the experimental DNA comprises promoter regions for each of the plurality of transcriptional regulators; (b) determining if any one of the plurality of transcriptional regulators are regulated by at least one of the plurality of transcriptional regulators; wherein a transcriptional regulatory network is identified if any one of the plurality of transcriptional regulators is regulated by at least one of the plurality of transcriptional regulators.

10

Specific embodiments of the methods for identifying regulatory networks described herein further comprise determining if any of the genes regulated by one of the plurality of transcriptional regulators is also a target of any of the other transcriptional regulators

15

The invention further provides algorithms for the identification of regulatory motifs, which may be used in conjunction with any of the methods provided herein, such as the methods for identifying the genes regulated by a transcriptional regulator. In a specific embodiment, two data matrices are created. The overall matrix D consists of binary entries D_{ij} , where a 1 indicates binding of regulator j to intergenic region i , a 0 indicates no binding event. The regulator matrix R is a subset of D , containing only the rows corresponding to the intergenic region assigned to each regulator, in the same order as the columns of regulators. The analyses may be performed using Matlab® software. The algorithms to find each motif are described as follows:

25

Autoregulatory motif: Find each non-zero entry on the diagonal of R .

Feedforward loop: For each master regulator (column of R), find non-zero entries, which correspond to regulators bound. For each master regulator / secondary regulator pair, find all rows in D bound by both regulators.

30

Multi-component loop: For each regulator (column of R), find the regulators to

which it binds. For each of these, find the regulators it binds. If any of these are the original regulator, you have a multi-component loop of two. For all others, find regulators to which they bind. If any of these are the original, you have a multicomponent loop of three. Repeat to find larger loops.

5

Single input module: Find the intergenic regions bound by only one regulator. That is, take the subset of rows of D such that the sum of each row is 1. Then for each regulator (column), find non-zero entries. Each set (greater than three intergenic regions) is a SIM.

10

Multi-input module: Find the intergenic regions bound by more than one regulator. That is, take the subset of rows of D such that the sum of each row is greater than 1. Then, for each row, find any other row bound by the same regulators. The collection of rows bound by the same regulators correspond to a MIM. Once a row is assigned to a MIM, remove it from further analysis.

15

Regulator chain: For each regulator (column of R), use a recursive algorithm to find chains of all lengths. That is, for each regulator whose promoter is bound by the regulator before it in the chain, find the regulator promoters to which it binds. Repeat until the chain ends. There are three possible ways to end a chain: a regulator that does not bind to the promoter of any other regulator, a regulator that binds to its own promoter, or one that binds to the promoter of another regulator earlier in the chain.

20

In one preferred embodiment of any of the methods described herein such as the methods for identifying regulatory networks, the experimental DNA in the microarray comprises promoter regions from additional transcriptional regulators or from genes suspected to encode transcriptional regulators. Such microarray enables one skilled in the art to identify the components of a regulatory pathway. For example, starting with one transcriptional regulator, a subset of the genes it regulates are identified using any method, such as those described herein. If one identified gene is itself a second transcriptional regulator or is suspected to encode a transcriptional regulator, then the subset of genes the second transcriptional regulator regulates is identified, and so on.

25
30

Furthermore, the subset of genes that the first and second transcriptional regulators regulate can be compared to determine if any genes are found in both subsets. If so, then a feed-forward motif, a unit of a regulatory network, has been identified.

Likewise, if the second transcriptional regulator is found to regulate the first one, then a
5 feedback loop has been identified.

4. Development of a Therapeutic to Treat or Prevent Disorders

One aspect of the invention provides methods of identifying targets for the development of therapeutics. One aspect of the invention provides a method of
10 identifying at least one target gene for the development of a therapeutic to treat or prevent a disorder in a subject, wherein at least one form of the disorder is caused by an altered activity in a transcriptional regulator or in a suspected transcriptional regulator, the method comprising (a) identifying the genes regulated by the transcriptional regulator in a cell; (b) determining if the transcriptional regulator is a broad-acting
15 transcriptional regulator or a narrow-acting transcriptional regulator, wherein if the transcriptional regulator is a broad acting transcriptional regulator then the transcriptional regulator is a target gene for the development of a therapeutic, and wherein if the transcriptional regulator is a narrow acting transcriptional regulator then (i) determining if at least one gene regulated by the transcriptional regulator is likely
20 causative in the disorder, wherein a gene that is likely causative in the disorder is a target gene for the development of a therapeutic; and (ii) reiterating steps (a) and (b) for at least one gene that is regulated by the transcriptional regulator in the cell and that either (1) encodes a transcriptional regulator or (2) is suspected to encode a transcriptional regulator, with the modification that the transcriptional regulator of steps
25 (a) and (b) is said gene, thereby identifying at least one target gene for the development of a therapeutic to treat or prevent a disorder in the subject.

In some embodiments of the methods for identifying a target gene for the development of a therapeutic, the genes regulated by the transcriptional regulator in the
30 cell are identified using chromosome-wide location analysis, analysis of mRNA transcripts in a cell that expresses the transcriptional regulator, or by using any of the methods provided herein for the identification of the genes that are regulated by a

transcriptional regulator. Some methods may comprise the use of DNA microarray or DNA arrays, such as those described in Gabrielson et al., Obesity Research, 8(5), 374-384 (2000).

5 In some embodiments of the methods described herein for identifying a target gene for the development of a therapeutic, the transcriptional regulator is a master regulatory gene. In specific embodiments, the master regulatory gene is SOX1-18, OCT6, PAX3, Myocardin, GATA1-6, TCF1/HNF1A, HNF4A, HNF6, NGN3, C/EBP, FOXA1-3, IPF1, GATA, HNF3, NKX2.1, CDX, FTF/NR5A2,
10 C/EBPbeta, SCL1, SKIN1, or a member of the neurogenin, LK, LMO, SOX, OCT, PAX, GATA or MyoD family of transcription factors.

 In some embodiments of the methods described herein, the transcriptional regulator is PAX3, EGR-1, EGR-2, OCT6, a SOX family member, a GATA family
15 member, a PAX family member, an OCT family member, RFX5, WHN, GATA1, VDR, CRX, CBP, MeCP2, AML1, p53, PLZF, PML, Rb, WT1, NR3C2, GCCR, PPARgamma, SIM1, HNF1alpha, HNF1beta, HNF4alpha, PDX1, MAFA, FOXA2, or NEUROD1.

20 A transcriptional regulator whose altered activity can lead to disease might be expressed in multiple, or all tissues of an organism, such that any of multiple cell types may be used in identifying a therapeutic. In some embodiments of the methods described herein for identifying a target gene for the development of a therapeutic, the cell is derived from a tissue whose function is impaired in the disorder. For example, a
25 pancreatic cell may be used for diabetes, a cardiac muscle cells for myocardial infarction, or neurons for Alzheimer's disease.

 In specific embodiments of the methods described herein for identifying a target gene for the development of a therapeutic, the broad acting gene regulates at least about
30 1%, 2% or more preferably at least about 2.5% of the genes in the cell, and the narrow acting gene regulates less than about 1%, 2% or 2.5% of the genes in the cell.

In specific embodiments of the methods described herein, a gene is suspected to encode a transcriptional regulator if it shares at least about 30%, 40% or 50% amino acid sequence identity within at least the DNA binding domain of a transcriptional regulator. DNA binding domains and methods of performing nucleic acids and polypeptide sequence alignments are well-known in the art. Optimal alignment of sequences for comparison may be conducted by the local homology algorithm of Smith and Waterman, *Adv. Appl. Math.* 2: 482 (1981); by the homology alignment algorithm of Needleman and Wunsch, *J. Mol Biol.* 48: 443 (1970); by the search for similarity method of Pearson and Lipman, *Proc. Natl. Acad. Sci.* 8: 2444 (1988); by computerized implementations of these algorithms, including, but not limited to: CLUSTAL in the PC/Gene program by Intelligenetics, Mountain View, Calif., GAP, BESTFIT, BLAST, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group (GCG), 7 Science Dr., Madison, Wis., USA; the CLUSTAL program is well described by Higgins and Sharp, *Gene*, 73: 237-244, 1988; Higgins and Sharp, *CABIOS* :11-13, 1989; Corpet, et al., *Nucleic Acids Research*, 16:881-90,1988; Huang, et al., *Computer Applications in the Biosciences* 8:1-7,1992; and Pearson, et al., *Methods in Molecular Biology* 24:7-331,1994.

In some specific embodiments of the methods described herein for identifying a target gene for the development of a therapeutic, the gene regulated by the transcriptional regulator is said to be likely causative of the disorder if a mutation in said gene results in at least one phenotype or symptom associated with the disorder. In another specific embodiment, the gene regulated by the transcriptional regulator is said to be likely causative of the disorder when the gene encodes an enzyme or signaling molecule which functions in a pathway that is impaired in the disorder. For example, if the disease is type II diabetes, a disorder characterized by hyperglycemia, then a gene regulated by the transcriptional regulator which encodes a sugar transporter, an enzyme involved in catalyzing a step of glycolysis or gluconeogenesis, or a gene which regulates insulin production, secretion or signaling is said to be likely causative of the disorder. In another specific embodiment, the gene regulated by the transcriptional regulator is said to be likely causative of the disorder if a mutant allele of the gene is genetically linked to a "susceptibility locus" for at least one form of the disease. A

“susceptibility locus” for a particular disease is a sequence or gene locus implicated in the initiation or progression of the disease. The susceptibility locus can be, for example, a gene or a microsatellite repeat, as identified by a microsatellite marker, or can be identified by a defined single nucleotide polymorphism. Generally, susceptibility genes
5 implicated in specific diseases and their loci can be found in scientific publications, but may also be determined experimentally.

In some embodiments of the methods described herein for identifying a target gene for the development of a therapeutic, the altered activity in the transcriptional
10 regulator comprises at least one of the following: (a) an alteration in the binding affinity of the transcriptional regulator to DNA; (b) an alteration in the ability of the transcriptional regulator to bind to RNA polymerase, to an RNA polymerase holoenzyme, or to a second transcriptional regulator; (c) an alteration in the binding
15 affinity of the transcriptional regulator to a ligand; (d) an alteration in expression level or expression pattern of the transcriptional regulator; or (e) an alteration in an ability of the transcriptional regulator to form homomultimers or heteromultimers.

In some embodiments of the methods described herein, the cell comprises a mutant form of the transcriptional regulator. A preferred mutant form of the
20 transcriptional regulator is one that causes the disease to which the therapeutic is sought. Such embodiments are particularly preferred when a mutant transcriptional regulator which causes at least one form of the disease has an altered target specificity and thus the genes it regulates, or the extent to which it regulates their transcription, is altered when compared to the non-mutant form of the transcriptional regulator. Such
25 embodiments may allow the identification of therapeutic targets which might not have been identified if a wild-type form of the transcriptional regulator had been used. Mutations in the DNA binding domain, for example, may alter the target specificity of a transcriptional regulator by altering its affinity for various DNA binding sequences.

30 It is well-known to one skilled in the art that mutations in a transcriptional regulator may result in a hypomorphic, hypermorphic or neomorphic phenotype. Mutations may generally reduce the activity of a transcriptional regulator, may

generally increase its activity, or may confer novel properties, such as altering the range of targets or turning an activator into a repressor or vice versa. In any methods described herein, and in particular those for identifying the therapeutics, a cell expressing a transcriptional regulator having any of these changes in activity may be
5 used.

The methods described herein may be applied to any disorder for which a transcriptional regulator has been implicated. Examples of diseases and transcriptional regulators which cause them may be found in the scientific and medical literature by
10 one skilled in the art, including in Medical Genetics, L.V. Jorde et al., Elsevier Science 2003, and Principles of Internal Medicine, 15th edition, ed by Braunwald et al., McGraw-Hill, 2001; American Medical Association Complete Medical Encyclopedia (Random House, Incorporated, 2003); and The Mosby Medical Encyclopedia, ed by Glanze (Plume, 1991). In some embodiments, the disorder is characterized by
15 impaired function of at least one of the following: brain, spinal cord, heart, arteries, esophagus, stomach, small intestine, large intestine, liver, pancreas, lungs, kidney, urinary tract, ovaries, breasts, uterus, testis, penis, colon, prostate, bone, muscle, cartilage, thyroid gland, adrenal gland, pituitary, bone marrow, blood, thymus, spleen, lymph nodes, skin, eye, ear, nose, teeth or tongue.

20

In some embodiments of the methods described herein for identifying a target gene for the development of a therapeutic, the subject is a mammal. In preferred embodiments, the subject is a human. In some embodiments of the methods described herein for identifying a target gene for the development of a therapeutic, the therapeutic
25 comprises a small molecule drug, an antisense nucleic acid, an antibody, a peptide, a ligand, a fatty acid, a hormone or a metabolite.

Antisense nucleic acids acting by RNAi include oligonucleotides which specifically hybridize (e.g., bind) under cellular conditions with a gene sequence, such
30 as at the cellular mRNA and/or genomic DNA level, so as to inhibit expression of that gene, e.g., by inhibiting transcription and/or translation. The binding may be by conventional base pair complementarity, or, for example, in the case of binding to DNA

duplexes, through specific interactions in the major groove of the double helix. Preferred antisense nucleic acid comprise siRNA, shRNAs, or any other form of double stranded RNA molecule. Antisense nucleic acids may be chemically modified, such as to increase their in vivo stability.

5

RNAi is a process of sequence-specific post-transcriptional gene repression which can occur in eukaryotic cells. In general, this process involves degradation of an mRNA of a particular sequence induced by double-stranded RNA (dsRNA) that is homologous to that sequence. For example, the expression of a long dsRNA
10 corresponding to the sequence of a particular single-stranded mRNA (ss mRNA) will labilize that message, thereby "interfering" with expression of the corresponding gene. Accordingly, any selected gene may be repressed by introducing a dsRNA which corresponds to all or a substantial part of the mRNA for that gene. It appears that when a long dsRNA is expressed, it is initially processed by a ribonuclease III into shorter
15 dsRNA oligonucleotides of in some instances as few as 21 to 22 base pairs in length. Furthermore, RNAi may be effected by introduction or expression of relatively short homologous dsRNAs. dsRNAs shorter than about 30 bases pairs are preferred to effect gene repression by RNAi (see Hunter et al. (1975) J Biol Chem 250: 409-17; Manche et al. (1992) Mol Cell Biol 12: 5239-48; Minks et al. (1979) J Biol Chem 254: 10180-3;
20 and Elbashir et al. (2001) Nature 411: 494-8).

Antibodies include whole antibodies, e.g., of any isotype (IgG, IgA, IgM, IgE, etc.), and includes fragments thereof which are also specifically reactive with a vertebrate, e.g., mammalian, protein. Antibodies may be fragmented using conventional
25 techniques and the fragments screened for utility in the same manner as described above for whole antibodies. Thus, the term includes segments of proteolytically-cleaved or recombinantly-prepared portions of an antibody molecule that are capable of selectively reacting with a certain protein. Non-limiting examples of such proteolytic and/or recombinant fragments include Fab, F(ab')₂, Fab', Fv, and single chain
30 antibodies (scFv) containing a V[L] and/or V[H] domain joined by a peptide linker. The scFv's may be covalently or non-covalently linked to form antibodies having two or more binding sites. The subject invention includes polyclonal, monoclonal,

humanized, or other purified preparations of antibodies and recombinant antibodies.

Peptidomimetic include compounds containing peptide-like structural elements that is capable of mimicking the biological action (s) of a natural parent polypeptide.

5

Hormone include any one of a number of biochemical substances that are produced by a certain cell or tissue and that cause a specific biological change or activity to occur in another cell or tissue located elsewhere in the body.

10 Metabolites includes any substance produced by metabolism or by a metabolic process. "Metabolism", as used herein, refers to the various chemical reactions involved in the transformation of molecules or chemical compounds occurring in tissue and the cells therein.

15 Ligands include any substance which binds to a receptor protein. A ligand of a transcriptional regulator protein is a substance which binds to the regulator protein, such as estrogen binding to a nuclear hormone receptor. In a preferred embodiment, ligand binding of to a transcriptional regulator occurs with high affinity. The term ligand refers to substances including, but not limited to, a natural ligand, whether
20 isolated and/or purified, synthetic, and/or recombinant, a homolog of a natural ligand (e.g., from another mammal). The term ligand encompasses substances which are inhibitors or promoters of receptor activity, as well as substances which selectively bind receptors, but lack inhibitor or promoter activity.

25 Some aspects of the invention relate to the diagnosis of disease states. A "transcriptional fingerprint", or listing of the genes, and optionally to what extent, that are regulated by given a transcriptional regulator can be generated from healthy individuals and from those afflicted with a disorder. Comparison of the fingerprints between the two groups may define genes which are specific to one of the two groups,
30 and thus serve as diagnostic for the risk that a patient is at risk, or is afflicted, with the disorder. In one embodiment, the transcriptional fingerprint of HNF4a is used to diagnose type II diabetes. A biopsy of a subject's liver or pancreas may provide the

cells for such analysis.

In specific embodiments, the transcriptional fingerprint disease diagnosis analysis is applied to transcriptional regulators which are causative in a particular disease to diagnose the disease. This approach may be coupled to allelic genotyping of the transcriptional regulator gene in the subject. For example, genotyping of a subject's HNF4a may uncover a novel allele. By using "transcriptional fingerprint" of HNF4a in tissue from that patient, one skilled in the art may determine what effect that mutation has in HNF4a activity and thus diagnose type II diabetes.

10

5. Methods of Preventing/Treating Disease through Regulation of HNFs

Some aspects of the invention provide methods of treating or preventing disease by regulating transcriptional regulator activity, particularly that of the HNF family member. The invention provides a method of treating or preventing type II diabetes in a subject, comprising administering to the subject a therapeutically effective amount of an agent that increases the global transcriptional activity of HNF4alpha. U.S. Patent No. 5,849,485 describes methods and assays for the isolation of modulators of HNF-4a activity, hereby incorporated by reference.

The invention also provides a method of treating or preventing a disorder associated with low transcriptional activity of HNF4alpha in a subject, comprising administering to the subject a therapeutically effective amount of an agent that increases the global transcriptional activity of HNF4alpha. In a related aspect, the invention provides a method of treating or preventing a disorder associated with high transcriptional activity of HNF4alpha in a subject, comprising administering to the subject a therapeutically effective amount of an agent that decreases the global transcriptional activity of HNF4alpha.

Yet another related aspect of the invention provides a method of increasing the global transcriptional activity in a liver or a pancreatic cell comprising contacting the cell with an agent which increases the global transcriptional activity of HNF4alpha. Similarly, the invention provides a method of decreasing the global transcriptional

activity in a liver or a pancreatic cell comprising contacting the cell with an agent which decreases the global transcriptional activity of HNF4alpha.

Applicants have identified genes that are transcriptionally regulated by HNF-1a,
5 HNF4a and HNF6 in hepatocytes and pancreatic cells. Accordingly, the invention provides methods of regulating the expression level of any of these genes in a cell or in a subject by contacting the cell or administering to the subject and agent which modulates the expression level or transcriptional regulatory activity of HNF transcription factors.

10

The invention provides a method of regulating the expression level of any one of the genes in Figure 13 in a hepatocyte, the method comprising contacting the cell with an agent which regulates the transcriptional activity of HNF1alpha. Similarly, the invention also provides a method of regulating the expression level of any one of the
15 genes in Figure 14 in a pancreatic cell, the method comprising contacting the cell with an agent which regulates the transcriptional activity of HNF1alpha.

The invention also provides a method of regulating the expression level of any one of the genes in Figure 16 in a hepatocyte, the method comprising contacting the
20 cell with an agent which regulates the transcriptional activity of HNF6. Similarly, the invention provides a method of regulating the expression level of any one of the genes in Figure 17 in a pancreatic cell, the method comprising contacting the cell with an agent which regulates the transcriptional activity of HNF6.

The invention additionally provides a method of regulating the expression level
25 of any one of the genes in Figure 18 in a hepatocyte, the method comprising contacting the cell with an agent which regulates the transcriptional activity of HNF4alpha. Similarly, the invention provides a method of regulating the expression level of any one of the genes in Figure 19 in a pancreatic cell, the method comprising contacting the cell with an agent which regulates the transcriptional activity of HNF4alpha.

30

Agents which modulate the transcriptional activity of HNF-4a, or any other HNF family member, may be identified by screening compounds for their ability to

increase the expression level, the DNA binding activity or the transcriptional promoting activity of HNF4a . One assay format which can be used employs two genetic constructs. One is typically a plasmid that continuously expresses the transcriptional regulator of interest when transfected into an appropriate cell line. CV-1 cells are most often used. The second is a plasmid which expresses a reporter, e.g., luciferase under control of the transcriptional regulator. For example, if a compound which acts as a ligand for HNF-4 is to be evaluated, one of the plasmids would be a construct that results in expression of the HNF-4 receptor in an appropriate cell line, e.g., the CV-1 cells. The second would possess a promoter linked to the luciferase gene in which an HNF-4 response element is inserted. If the compound to be tested is an agonist for the HNF-4 receptor, the ligand will complex with the receptor and the resulting complex binds the response element and initiates transcription of the luciferase gene. In time the cells are lysed and a substrate for luciferase added. The resulting chemiluminescence is measured photometrically. Dose response curves are obtained and can be compared to the activity of known ligands. Other reporters than luciferase can be used including CAT and other enzymes.

Viral constructs can be used to introduce the gene for the receptor and the reporter. An usual viral vector is an adenovirus. For further details concerning this preferred assay, see U.S. Pat. No. 4,981,784 issued Jan. 1, 1991 hereby incorporated by reference, and Evans et al., WO88/03168 published on 5 May 1988, also incorporated by reference.

HNF-4a antagonists can be identified using this same basic "agonist" assay. A fixed amount of an antagonist is added to the cells with varying amounts of test compound to generate a dose response curve. If the compound is an antagonist, expression of luciferase is suppressed.

Additional methods for the isolation of agonists and antagonist of HNF transcription factors are described in U.S. Patent Nos. 6,187,533 and 5,620,887. Additional U.S. patents describing methods to identify agents that modulate the activity of transcription factors include 5,804,374, and 5,298,429, and U.S. Patent Publication

Nos. 2004/0033942A1 2003/0077664, 2003/0215829 and 2003/0039980. Any of the methods described herein may be easily adapted to identify agonists or antagonists of any one of the HNF transcriptional factors. U.S. Patent No. 6,303,653 describes modulators of HNF-4 activity.

5

Agonists and antagonists of HNF4a can also be designed based on the known crystal structure of HNF4a complexed with an endogenous fatty acid ligand (Dhe-Paganon, J. Biol. Chem. 277(41), 37973-37976). U.S. Patent Publication No. 2002/0072587 describes methods of identifying agonists of an estrogen receptor, a nuclear receptor like the HNF proteins, based on its crystal structure. Such methods may easily be applied to HNF-1a, HNF-4a and HNF6 by one skilled in the art. Additional examples of rational drug design based on the structure of a protein may be found in U.S. Patent or Publication Nos. 6,236,946, 6,684,162, 2004/0014153, 2003/0124699, 20030077628, 2002/0151028, 2002/0072587 and 2003/0211588.

15

6. Therapeutics

In one aspect, the invention provides methods of treating disease in a subject comprising the administration of a composition comprising a therapeutic agent. "Therapeutic agent" or "therapeutic" refers to an agent capable of having a desired biological effect on a host. Chemotherapeutic and genotoxic agents are examples of therapeutic agents that are generally known to be chemical in origin, as opposed to biological, or cause a therapeutic effect by a particular mechanism of action, respectively. Examples of therapeutic agents of biological origin include growth factors, hormones, and cytokines. A variety of therapeutic agents are known in the art and may be identified by their effects. Certain therapeutic agents are capable of regulating cell proliferation and differentiation. Examples include chemotherapeutic nucleotides, drugs, hormones, non-specific (non-antibody) proteins, oligonucleotides (e.g., antisense oligonucleotides that bind to a target nucleic acid sequence (e.g., mRNA sequence)), peptides, and peptidomimetics.

30

In one embodiment, the compositions are pharmaceutical compositions. Pharmaceutical compositions for use in accordance with the present invention may be

formulated in conventional manner using one or more physiologically acceptable carriers or excipients. Thus, the compounds and their physiologically acceptable salts and solvates may be formulated for administration by, for example, by aerosol, intravenous, oral or topical route. The administration may comprise intralesional, intraperitoneal, subcutaneous, intramuscular or intravenous injection; infusion; liposome-mediated delivery; topical, intrathecal, gingival pocket, per rectum, intrabronchial, nasal, transmucosal, intestinal, oral, ocular or otic delivery.

An exemplary composition of the invention comprises an compound capable of modulating the expression or activity of a transcriptional regulator with a delivery system, such as a liposome system, and optionally including an acceptable excipient. In a preferred embodiment, the composition is formulated for injection.

Techniques and formulations generally may be found in Remington's Pharmaceutical Sciences, Meade Publishing Co., Easton, PA. For systemic administration, injection is preferred, including intramuscular, intravenous, intraperitoneal, and subcutaneous. For injection, the compounds of the invention can be formulated in liquid solutions, preferably in physiologically compatible buffers such as Hank's solution or Ringer's solution. In addition, the compounds may be formulated in solid form and redissolved or suspended immediately prior to use. Lyophilized forms are also included.

For oral administration, the pharmaceutical compositions may take the form of, for example, tablets or capsules prepared by conventional means with pharmaceutically acceptable excipients such as binding agents (e.g., pregelatinised maize starch, polyvinylpyrrolidone or hydroxypropyl methylcellulose); fillers (e.g., lactose, microcrystalline cellulose or calcium hydrogen phosphate); lubricants (e.g., magnesium stearate, talc or silica); disintegrants (e.g., potato starch or sodium starch glycolate); or wetting agents (e.g., sodium lauryl sulphate). The tablets may be coated by methods well known in the art. Liquid preparations for oral administration may take the form of, for example, solutions, syrups or suspensions, or they may be presented as a dry product for constitution with water or other suitable vehicle before use. Such liquid

preparations may be prepared by conventional means with pharmaceutically acceptable additives such as suspending agents (e.g., sorbitol syrup, cellulose derivatives or hydrogenated edible fats); emulsifying agents (e.g., lecithin or acacia); non-aqueous vehicles (e.g., ationd oil, oily esters, ethyl alcohol or fractionated vegetable oils); and
5 preservatives (e.g., methyl or propyl-p-hydroxybenzoates or sorbic acid). The preparations may also contain buffer salts, flavoring, coloring and sweetening agents as appropriate.

Preparations for oral administration may be suitably formulated to give
10 controlled release of the active compound. For buccal administration the compositions may take the form of tablets or lozenges formulated in conventional manner. For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebuliser, with the use of a suitable propellant, e.g.,
15 dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of e.g., gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose
20 or starch.

The compounds may be formulated for parenteral administration by injection, e.g., by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, e.g., in ampoules or in multi-dose containers, with an
25 added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents. Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, e.g., sterile pyrogen-free water, before use.

30

The compounds may also be formulated in rectal compositions such as suppositories or retention enemas, e.g., containing conventional suppository bases such

as cocoa butter or other glycerides.

In addition to the formulations described previously, the compounds may also be formulated as a depot preparation. Such long acting formulations may be administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, the compounds may be formulated with suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration bile salts and fusidic acid derivatives. In addition, detergents may be used to facilitate permeation. Transmucosal administration may be through nasal sprays or using suppositories. For topical administration, the oligomers of the invention are formulated into ointments, salves, gels, or creams as generally known in the art. A wash solution can be used locally to treat an injury or inflammation to accelerate healing.

The compositions may, if desired, be presented in a pack or dispenser device which may contain one or more unit dosage forms containing the active ingredient. The pack may for example comprise metal or plastic foil, such as a blister pack. The pack or dispenser device may be accompanied by instructions for administration.

For therapies involving the administration of nucleic acids, the oligomers of the invention can be formulated for a variety of modes of administration, including systemic and topical or localized administration. Techniques and formulations generally may be found in Remington's Pharmaceutical Sciences, Meade Publishing Co., Easton, PA. For systemic administration, injection is preferred, including intramuscular, intravenous, intraperitoneal, intranodal, and subcutaneous for injection, the oligomers of the invention can be formulated in liquid solutions, preferably in

physiologically compatible buffers such as Hank's solution or Ringer's solution. In addition, the oligomers may be formulated in solid form and redissolved or suspended immediately prior to use. Lyophilized forms are also included.

5 Systemic administration can also be by transmucosal or transdermal means, or the compounds can be administered orally. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration bile salts and fusidic acid derivatives. In addition,
10 detergents may be used to facilitate permeation. Transmucosal administration may be through nasal sprays or using suppositories. For oral administration, the oligomers are formulated into conventional oral administration forms such as capsules, tablets, and tonics. For topical administration, oligomers may be formulated into ointments, salves, gels, or creams as generally known in the art.

15 Toxicity and therapeutic efficacy of the agents and compositions of the present invention can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., for determining the LD₅₀ (the dose lethal to 50% of the population) and the ED₅₀ (the dose therapeutically effective in 50% of the population).
20 The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD₅₀/ED₅₀. Compounds which exhibit large therapeutic induces are preferred. While compounds that exhibit toxic side effects may be used, care should be taken to design a delivery system that targets such compounds to the site of affected tissue in order to minimize potential damage to uninfected cells and,
25 thereby, reduce side effects.

 The data obtained from the cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED₅₀ with little
30 or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially

from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the IC₅₀ (i.e., the concentration of the test compound which achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by high performance liquid chromatography.

In one embodiment of the methods described herein, the effective amount of the agent is between about 1mg and about 50mg per kg body weight of the subject. In one embodiment, the effective amount of the agent is between about 2mg and about 40mg per kg body weight of the subject. In one embodiment, the effective amount of the agent is between about 3mg and about 30mg per kg body weight of the subject. In one embodiment, the effective amount of the agent is between about 4mg and about 20mg per kg body weight of the subject. In one embodiment, the effective amount of the agent is between about 5mg and about 10mg per kg body weight of the subject.

In one embodiment of the methods described herein, the agent is administered at least once per day. In one embodiment, the agent is administered daily. In one embodiment, the agent is administered every other day. In one embodiment, the agent is administered every 6 to 8 days. In one embodiment, the agent is administered weekly.

As for the amount of the compound and/or agent for administration to the subject, one skilled in the art would know how to determine the appropriate amount. As used herein, a dose or amount would be one in sufficient quantities to either inhibit the disorder, treat the disorder, treat the subject or prevent the subject from becoming afflicted with the disorder. This amount may be considered an effective amount. A person of ordinary skill in the art can perform simple titration experiments to determine what amount is required to treat the subject. The dose of the composition of the invention will vary depending on the subject and upon the particular route of administration used. In one embodiment, the dosage can range from about 0.1 to about 100,000 ug/kg body weight of the subject. Based upon the composition, the dose can be

delivered continuously, such as by continuous pump, or at periodic intervals. For example, on one or more separate occasions. Desired time intervals of multiple doses of a particular composition can be determined without undue experimentation by one skilled in the art.

5

The effective amount may be based upon, among other things, the size of the compound, the biodegradability of the compound, the bioactivity of the compound and the bioavailability of the compound. If the compound does not degrade quickly, is bioavailable and highly active, a smaller amount will be required to be effective. The effective amount will be known to one of skill in the art; it will also be dependent upon the form of the compound, the size of the compound and the bioactivity of the compound. One of skill in the art could routinely perform empirical activity tests for a compound to determine the bioactivity in bioassays and thus determine the effective amount. In one embodiment of the above methods, the effective amount of the compound comprises from about 1.0 ng/kg to about 100 mg/kg body weight of the subject. In another embodiment of the above methods, the effective amount of the compound comprises from about 100 ng/kg to about 50 mg/kg body weight of the subject. In another embodiment of the above methods, the effective amount of the compound comprises from about 1 ug/kg to about 10 mg/kg body weight of the subject. In another embodiment of the above methods, the effective amount of the compound comprises from about 100 ug/kg to about 1 mg/kg body weight of the subject.

As for when the compound, compositions and/or agent is to be administered, one skilled in the art can determine when to administer such compound and/or agent. The administration may be constant for a certain period of time or periodic and at specific intervals. The compound may be delivered hourly, daily, weekly, monthly, yearly (e.g. in a time release form) or as a one time delivery. The delivery may be continuous delivery for a period of time, e.g. intravenous delivery. In one embodiment of the methods described herein, the agent is administered at least once per day. In one embodiment of the methods described herein, the agent is administered daily. In one embodiment of the methods described herein, the agent is administered every other day. In one embodiment of the methods described herein, the agent is administered every 6

to 8 days. In one embodiment of the methods described herein, the agent is administered weekly.

5 **EXEMPLIFICATION**

 The invention now being generally described, it will be more readily understood by reference to the following examples, which are included merely for purposes of illustration of certain aspects and embodiments of the present invention, and are not intended to limit the invention, as one skilled in the art would recognize from the teachings hereinabove and the following examples, that other DNA microarrays, transcriptional regulators, cell types, antibodies, ChIP conditions, or data analysis methods, all without limitation, can be employed, without departing from the scope of the invention as claimed.

 The practice of the present invention will employ, where appropriate and unless otherwise indicated, conventional techniques of cell biology, cell culture, molecular biology, transgenic biology, microbiology, virology, recombinant DNA, and immunology, which are within the skill of the art. Such techniques are described in the literature. See, for example, Molecular Cloning: A Laboratory Manual, 3rd Ed., ed. by Sambrook and Russell (Cold Spring Harbor Laboratory Press: 2001); the treatise, Methods In Enzymology (Academic Press, Inc., N.Y.); Using Antibodies, Second Edition by Harlow and Lane, Cold Spring Harbor Press, New York, 1999; Current Protocols in Cell Biology, ed. by Bonifacino, Dasso, Lippincott-Schwartz, Harford, and Yamada, John Wiley and Sons, Inc., New York, 1999; and PCR Protocols, ed. by Bartlett et al., Humana Press, 2003.

 Various publications, patents, and patent publications are cited throughout this application the contents of which are incorporated herein by reference in their entirety.

30 Experimental procedures

 The following procedures were followed in performing the experiments below:

Genome-scale Location Analysis

The protocol described here was adapted from Ren 2001. Briefly, cells are fixed with 1% final concentration formaldehyde for 10-20 minutes at room temperature, harvested and rinsed with 1x PBS. The resultant cell pellet is sonicated, and DNA fragments that are crosslinked to a protein of interest are enriched by immunoprecipitation with a factor specific antibody. After reversal of the crosslinking, the enriched DNA is amplified using ligation-mediated PCR (LM-PCR), and then fluorescently labeled using high concentration Klenow polymerase and a dNTP-fluorophore. A sample of DNA that has not been enriched by immunoprecipitation is subjected to LM-PCR and labeled with a different fluorophore. Both IP-enriched and unenriched pools of labeled DNA are hybridized to a single DNA microarray containing 13,000 human intergenic regions (see below for description of DNA microarray and binding site determination). For hepatocyte experiments, 2.5×10^7 hepatocytes were typically used per chromatin immunoprecipitation. These hepatocytes were isolated by standard liver perfusion techniques, immediately crosslinked with 1% formaldehyde solution, rinsed, and flash frozen. Islet preparations were treated with formaldehyde between 1 hour and 5 days after isolation from pancreata. A minimum of 30,000 viable islet equivalents (approximately 2×10^7 beta cells) were fixed and handled as described above. Typical islet purity for three experiments described here was >70% islets with >80% viability. HNF4a, HNF6, and RNA polymerase II produced high quality results with as few as 30,000 islet equivalents. HNF1a ChIP required significantly more material, typically 80,000 islets, to produce results with somewhat lower enrichment ratios than the results obtained with hepatocytes.

Human 13K DNA Microarray

It would be ideal to have a DNA microarray that contains the entire human genome sequence, but technical limitations and cost led applicants to select the most relevant portion of the genome for inclusion in this microarray. Because a significant percentage of transcriptional binding sites in proximal promoters are within 1 kb of transcription start sites, applicants designed primers to amplify these genomic regions for printing onto a promoter array. Applicants selected 15000 cDNAs from the NCBI RefSeq database, and mapped them to NCBI Build 22 (April 2001) of the human

genome using BLAST. Where multiple splice variants had been described, applicants used the most upstream site, and verified the 5'-end by alignment with the Database of Transcriptional Start Sites (<http://elmo.ims.utokyo.ac.jp/dbtss/>). Sequences to be amplified were extracted from the genomic region -750 bp to +250 bp relative to this transcriptional start site. To control for nonspecific binding, 9 amplified regions derived from long Arabidopsis open reading frames were included on the array. As a further negative control and for use in data normalization, applicants chose 158 ORF regions within long exons of human genes for amplification. To prepare the DNA content of the arrays, the program Primer3

10 (http://www.genome.wi.mit.edu/genome_software/other/primer3.html) was used to design primers using the sequences described above. PCRs were performed on these primer set using standard conditions, except for the presence of 1 M betaine in all PCR reactions. Betaine was empirically observed to increase the success rate of the amplification reactions.

15

Of the 13,000 PCR pairs, 70% gave a strong band of the appropriate size, as verified on 2% agarose gels. Applicants have noted, however, that PCR products undetectable by agarose EtBr gel analysis can give valid positive signals when concentrated and printed on the DNA arrays. PCR quality evaluations were performed on the BRIDNAsuite of programs from the Biotechnology Research Institute of the National Research Council of Canada (<http://www.irb-bri.cnrc-nrc.gc.ca/>). PCR products were recovered from the reaction mixture by ammonium acetate/isopropanol precipitation and resuspended into 3x SSC with 1.5 M betaine to minimize evaporation and improve spot quality. Applicants printed amplified products onto GAPS-coated

25 glass slides (Corning) using a Cartesian PixSys 5500 arrayer. The quality of the arrays was determined on a batch-wise basis by hybridization with sequence neutral oligonucleotides covalently linked to Cy3 or Cy5, followed by calculation of usable percentage of spots, combined with direct visual inspection of the quality of the chip. The Hu13K array was remapped post-production using two independent methods. First,

30 applicants performed electronic PCR on the primer sets against the August 2003 final release of the completed human genome. Second, applicants BLASTed the sequence used to extract primers for amplification against the August 2003 final release of the

human genome. The dataset downloadable from the supporting website reports the location of each arrayed promoter relative to the transcriptional start site.

Data Quality Control

5 1. ChIP Hybridization Quality Control

The raw data generated from each array experiment was subjected to multiple levels of quality control. First, each scan was examined visually as it was being performed. Samples on microarrays with gross defects (e.g. scratches, smeared spots) were repeated whenever possible. Applicants also determined that no reliable signal
10 was produced from control spots containing *Arabidopsis* DNA.

2. Binding Site Determination and Error Model

Scanned images were analyzed using GenePix (v3.1 or v4.0), to obtain background subtracted intensity values. Each spot is bound by both IP-enriched and
15 unenriched DNA, which are labeled with different fluorophores. Consequently, each spot yields fluorescence intensity information in two channels, corresponding to immunoprecipitated DNA and genomic DNA. To account for background hybridization to slides, the median intensity of a set of control blank spots was subtracted for site-specific transcription factors (e.g. HNF1a), and the median intensity for a set of control
20 ORF spots was subtracted for broadly acting DNA binding proteins (e.g. RNA Pol II, HNF4a). To correct for different amounts of genomic and immunoprecipitated DNA hybridized to the microarray, the median intensity value of the IP-enriched DNA channel was divided by the median of the genomic DNA channel, and this normalization factor was applied to each intensity in the genomic DNA channel. Next,
25 applicants calculated the log of the ratio of intensity in the IP-enriched channel to intensity in the genomic DNA channel for each intergenic region across the entire set of hybridization experiments. Adjusted intensity values for the IP-enriched channel were calculated from these ratios. A whole-chip error model (Hughes 2000; Lee 2002) was then used to calculate confidence values for each spot on each microarray, and to
30 combine data for the replicates of each experiment to obtain a final average ratio and confidence for each promoter region. Genes were included in the set of 'bound' genes if the binding P-value in the error model was < 0.001 or enrichment was at least 2-fold

in the immunoprecipitation.

Confirmation of Predicted Binding

The accuracy of genome-wide location data reported here has been assessed
5 using several approaches.

1. Estimation of False Positive Rates Using Conventional ChIP Experiments

Conventional, independent ChIP experiments conducted in our laboratory at a
gene specific level have confirmed over 100 binding interactions identified by location
10 analysis data involving 6 different regulators (see
<http://web.wi.mit.edu/young/pancregulators>). These results suggest that our empirical
rate of false positives is at most 16%. This rate is somewhat higher than that found for a
large scale survey of yeast transcription factors (Lee 2002), which probably reflects the
greater complexity of the human genome. Figures 9 and 10 show typical verification
15 ChIP experiments for HNF4a and HNF1a, respectively, in hepatocytes.

2. Comparison with Previous Literature

Applicants found no previous studies of the genomic targets of transcriptional
regulators in primary human tissue. However, a large number of HNF1a and HNF4a
targets have been identified in model organisms and human carcinoma (mostly
20 hepatoma) cell lines; these targets are summarized in Figure 14. For example, genome-
scale location analysis identified 30 of the 68 hepatocyte genes which were both
previously suggested to be targets of HNF4a, and included on the 13K DNA array.
Similarly, genome-scale location analysis identified 21 of the 81 hepatocyte genes
which were both previously suggested to be targets of HNF4a, and included on the 13K
25 DNA array. Discrepancies between the targets reported here and targets reported in the
literature may result from a number of factors, which include, but are not limited to: (1)
the limitations of using a 1 kb promoter fragment to probe the binding of a transcription
factor, (2) the stringency of our threshold criteria, (3) the differences between the
regulatory network in model organisms and/or cell lines, and the regulatory network in
30 primary human tissue, (4) differences between indirect technologies in the literature
(i.e. gel-shift and transient transfections) and genome-scale location analysis, (5) tissue
isolation effects, among others. A more comprehensive discussion can be found at

<http://web.wi.mit.edu/young/pancregulators>

Regulatory Motifs Derived from Binding Data

In order to discover network motifs, two data matrices were created. The overall matrix D consists of binary entries D_{ij} , where a 1 indicates binding of regulator j to intergenic region i , a 0 indicates no binding event. The regulator matrix R is a subset of D, containing only the rows corresponding to the intergenic region assigned to each regulator, in the same order as the columns of regulators. All analyses were performed in Matlab. The algorithms used to find each motif are described below.

Autoregulatory motif: Find each non-zero entry on the diagonal of R. Feedforward loop: For each master regulator (column of R), find non-zero entries, which correspond to regulators bound. For each master regulator / secondary regulator pair, find all rows in D bound by both regulators. Multi-component loop: For each regulator (column of R), find the regulators to which it binds. For each of these, find the regulators it binds. If any of these are the original regulator, you have a multi-component loop of two. For all others, find regulators to which they bind. If any of these are the original, you have a multicomponent loop of three. Repeat to find larger loops. Single input module: Find the intergenic regions bound by only one regulator. That is, take the subset of rows of D such that the sum of each row is 1. Then for each regulator (column), find non-zero entries. Each set (greater than three intergenic regions) is a SIM. Multi-input module: Find the intergenic regions bound by more than one regulator. That is, take the subset of rows of D such that the sum of each row is greater than 1. Then, for each row, find any other row bound by the same regulators. The collection of rows bound by the same regulators correspond to a MIM. Once a row is assigned to a MIM, remove it from further analysis. Regulator chain: For each regulator (column of R), use a recursive algorithm to find chains of all lengths. That is, for each regulator whose promoter is bound by the regulator before it in the chain, find the regulator promoters to which it binds. Repeat until the chain ends. There are three possible ways to end a chain: a regulator that does not bind to the promoter of any other regulator, a regulator that binds to its own promoter, or one that binds to the promoter of another regulator earlier in the chain.

Example 1

The liver and pancreas have long been the subject of studies to understand how organs develop and are regulated at the transcriptional level (8-12). The transcriptional regulators HNF1 α (a homeodomain protein), HNF4 α (a nuclear receptor) and HNF6 (a member of the onecut family) operate cooperatively in a connected network in the liver, but less is known about the structure of this regulatory network in human pancreatic islets. All three transcriptional regulators are required for normal function of liver and pancreatic islets (13-18). Mutations in HNF1 α and HNF4 α are the causes of the type 3 and type 1 forms of maturity-onset diabetes of the young (MODY3 and MODY1), a genetic disorder of the insulin-secreting pancreatic beta cells characterized by onset of diabetes mellitus before 25 years of age and an autosomal dominant pattern of inheritance (19).

Applicants hypothesized that genome-scale analysis of the pancreatic islet genes whose expression is regulated by these transcription factors in normal beta cells could provide insights into the molecular basis of the abnormal beta cell function that characterizes MODY. Applicants have identified the genes occupied by the transcription factors HNF1 α , HNF4 α , and HNF6 in pancreatic islets. The genes transcribed in each tissue were identified by determining the genomic occupancy of RNA polymerase II. Applicants used this information to begin to map the transcriptional regulatory circuitry in these tissues.

Applicants first used genome-scale location analysis (20) to identify the promoters bound by HNF1 α in human hepatocytes and pancreatic islets isolated from tissue donors (Fig 1A). For each tissue, HNF1 α -DNA complexes were enriched by chromatin immunoprecipitation in three separate experiments. Applicants constructed a custom DNA microarray containing portions of promoter regions of 13,000 human genes (Hu13K array). Applicants targeted the region spanning 700 bp upstream and 200 bp downstream of transcription start sites for the genes whose start sites are best characterized based on National Center for Biotechnology Information annotation (20). Although many enhancers are present at more distant locations, most known

transcription factor binding site sequences occur within these start-site proximal regions of promoters.

The results of these genome location experiments revealed that HNF1 α is
5 bound to at least 222 target genes in hepatocytes, representing 1.6% of the genes on the
Hu13K array (Figure 11) (20). This result was verified with independent, conventional
chromatin immunoprecipitation experiments, which suggest that the frequency of false
positives in genome-scale location data with gene-specific regulators is no more than
16% when our threshold criteria were used (20). The genes applicants found to be
10 occupied by HNF1 α in primary human hepatocytes encode products whose functions
represent a significant cross-section of hepatocyte biochemistry. The results confirm
that HNF1 α contributes to the transcriptional regulation of many of the central rate-
limiting steps in gluconeogenesis and associated pathways. HNF1 α also binds to genes
whose products are central to normal hepatic function, including carbohydrate synthesis
15 and storage, lipid metabolism (synthesis of cholesterol and apolipoproteins),
detoxification (synthesis of cytochrome P450s) and synthesis of serum proteins
(albumin, complements and coagulation factors).

Applicants next identified HNF1 α target genes in human pancreatic islets
20 (Figure 11) (20). HNF1 α occupied the promoter regions of 106 genes (0.8% of the
Hu13K array promoters) in islets, 30% of which were also bound by HNF1 α in
hepatocytes (Figure 1B). In islets, fewer chaperones and enzymes are bound by HNF1 α
than in hepatocytes, and the receptors and signal transduction machinery regulated by
HNF1 α vary between the two tissues.

25

HNF1 α has been previously implicated in the regulation of many genes in
hepatocytes and islets (13, 16, 20 [Figure 15]). The direct genome binding data
reported here confirmed many, but not all, of these genes. The difference may be due,
at least in part, to our stringent criteria for binding in the genome-scale data, which
30 enhances our confidence in the direct target genes identified by location analysis, but
likely underestimates the actual number of targets in vivo. Furthermore, although the

proximal promoter regions printed on the array contain a significant number of transcription factor binding sequences, many genes are also regulated by more distal promoter elements and enhancers that are not present on the Hu13K array.

5 Applicants also identified the promoters bound by HNF6 in human hepatocytes and pancreatic islets using genome-scale location analysis (Fig 1B; Figures 16 and 17) (20). HNF6 was bound to at least 222 genes in hepatocytes and 189 genes in pancreatic islets, representing 1.7% and 1.4% of the promoters on the array, respectively. Approximately half of the promoters occupied by HNF6 were common to the two
10 tissues, and included a number of important cell cycle regulators such as CDK2 (20).

Genome-scale location analysis revealed surprising results for HNF4 α in hepatocytes and pancreatic islets (Fig 1B). The number of genes enriched in HNF4 α chromatin immunoprecipitations was much larger than observed with typical site-
15 specific regulators. HNF4 α was bound to approximately 12% of the genes represented on the Hu13K DNA microarray in hepatocytes and 11% in pancreatic islets. No other transcription factor applicants have profiled in human cells has been observed to bind more than 2.5% of the promoter regions represented on the 13K array.

Six independent lines of evidence indicate that the HNF4 α results are not due to
20 poor antibody specificity or errors in the microarray analysis, and support the view that HNF4 α is associated with an unusually large number of promoters in hepatocytes and pancreatic islets (20). First, essentially identical results were obtained with two different antibodies that recognize different portions of HNF4 α . Second, Western blots showed that the HNF4 α antibodies are highly specific. Third, applicants verified
25 binding at over 50 randomly selected targets of HNF4 α in hepatocytes by conventional gene-specific chromatin immunoprecipitation. Fourth, when antibodies against HNF4 α were used for ChIP in control experiments with Jurkat, U937, and BJT cells (which do not express HNF4 α), no more than 17 promoters were identified in each cell line by our criteria, which is well within the noise inherent in this system. Fifth, when pre-
30 immune antibodies from rabbit and goat (the two different anti-HNF4 α antibodies came from rabbit and goat) were used in control experiments in hepatocytes, the

number of targets identified was within the noise. Finally, if the HNF4 α results are correct, then applicants would expect that the set of promoters bound by HNF4 α should be largely a subset of those bound by RNA polymerase II in each tissue; applicants found that this is the case (see below). Applicants conclude that HNF4 α is a widely
5 acting transcription factor in these tissues, consistent with the observation that it is an unusually abundant, constitutively active transcription factor (11).

Applicants next identified the genes represented on the Hu13K microarray that are actively transcribed in hepatocytes and pancreatic islets, so the fraction of actively
10 transcribed genes that are bound by HNF4 α could be determined (Fig 2C). It is difficult to determine accurately the transcriptome of these tissues by profiling transcript levels with DNA microarrays. Transcript profiling requires a reference RNA population against which a tissue RNA population can be compared, and there are limitations to generating appropriate reference RNA. To circumvent this limitation,
15 applicants exploited the fact that RNA polymerase II occupies the set of protein-coding genes that are actively transcribed in eukaryotic cells. Location analysis with RNA polymerase II antibodies can identify these actively transcribed genes (7, 21). Applicants found that 23% of the genes on the Hu13K array (2984 genes) were bound by RNA polymerase II in hepatocytes, and 19% (2426 genes) were bound by RNA
20 polymerase II in islets (20). The sets of genes occupied by RNA polymerase II in hepatocytes and islets overlapped substantially (81% overlap, relative to islets), consistent with the relatedness of the two tissues (22). As expected, the majority of genes occupied by HNF4 α in hepatocytes and pancreatic islets (80% and 73%, respectively) were also occupied by RNA polymerase II. Remarkably, of the genes
25 occupied by RNA polymerase II, 42% (1262/2984) were bound by HNF4 α in hepatocytes and 43% (1047/2426) were bound by HNF4 α in islets (Fig 1C). By comparison, only 6% and 2% of RNA polymerase II enriched promoters were also bound by HNF1 α in hepatocytes and islets, respectively.

30 Previous studies indicate that HNF1 α , HNF4 α , and HNF6 are at the center of a network of transcription factors that cooperatively regulate numerous developmental and metabolic functions in hepatocytes and islets (9, 13, 15, 17). Our systematic

analysis of the direct in vivo targets of these factors significantly expands our understanding of the regulatory network in primary human tissues (Fig 2A). A comparison of the regulatory network in these two tissues reveals that HNF1 α , HNF4 α , and HNF6 occupy the promoters of genes encoding a large population of transcription factors and cofactors in the two tissues (20). The precise set of transcription factor genes occupied by HNF1 α , HNF4 α , and HNF6, and the extent to which they are co-occupied by the HNF regulators, differed substantially between these two tissues.

The transcription factor binding data was used to identify regulatory network motifs, simple units of transcriptional regulatory network architecture that suggest mechanistic models (Fig 2B) (4, 23). Our data confirm previous reports that HNF1 α and HNF4 α occupy one another's promoters in both hepatocytes and islets, forming a multi-component loop (24-26). Multicomponent loops provide the capacity for feedback control and produce bistable systems that can switch between two alternate states (23). It has been suggested that the multicomponent loop present between HNF1 α and HNF4 α is responsible for stabilization of the terminal phenotype in pancreatic beta cells (26). Applicants also found that HNF6 serves as a master regulator for feedforward motifs in hepatocytes and pancreatic islets involving over 80 genes in each tissue (Figures 20 and 22). For example, in hepatocytes, HNF6 binds the HNF4 α 7 promoter, and HNF6 and HNF4 α together bind *PCK1*, which encodes phosphoenolpyruvate carboxykinase, an enzyme key to gluconeogenesis (Fig 2B). A feedforward loop can act as a switch designed to be sensitive to sustained, rather than transient, inputs (23). HNF1 α , HNF4 α and HNF6 were also found to form multi-input motifs by collectively binding to sets of genes in hepatocytes and islets. This regulatory motif suggests coordination of gene expression through multiple input signals. Applicants also found that HNF6, HNF4 α , and HNF1 α form a regulator chain motif with THRA (NR1D1); regulator chain motifs represent the simplest circuit logic for ordering transcriptional events in a temporal sequence (4, 23). Additional examples of these regulatory motifs can be found in Figures 20 and 23 (20). Figures 20-24, panels A and B, show transcriptional regulators occupied by HNF transcription factors and their regulatory loops. Figures 4-10 show additional controls and data generated by the experiments described herein.

Our results suggest that the nuclear hormone receptor HNF4 α contributes to regulation of a large fraction of the liver and pancreatic islet transcriptomes by binding directly to almost half of the actively transcribed genes. This likely explains why

5 HNF4 α is crucial for development and proper function of these tissues (12-15, 17, 18). Perhaps most importantly, our results suggest a mechanistic explanation for the recent discovery that polymorphisms in the islet-specific P2 promoter for the splice variant HNF4 α 7 can greatly increase the risk of type II diabetes (27-30). Applicants found that multiple HNF factors bind directly to the P2 promoter in primary, healthy human islets.

10 Alterations in the binding sites for these factors could cause misregulation of HNF4 α expression and thus its downstream targets, leading to beta cell malfunction and diabetes.

References for Experimental Section:

- 15 1. Roeder, R.G. *Cold Spring Harb Symp Quant Biol* 63, 201 (1998).
2. T.I. Lee, R.A. Young. *Annu Rev Genet* 34, 77 (2000).
3. G. Orphanides, D. Reinberg. *Cell* 108, 439 (2002).
4. T.I. Lee, et al. *Science* 298, 799 (2002).
5. B. Ren, et al. *Genes Dev* 16, 245 (2002).
- 20 6. A.S. Weinmann, et al. *Genes Dev* 16, 235 (2002).
7. Z. Li, et al. *Proc Natl Acad Sci U S A* 100, 8164 (2003).
8. E. Lai, J.E. Darnell, Jr. *Trends Biochem Sci* 16, 427 (1991).
9. C.J. Kuo, et al. *Nature* 355, 457 (1992).
10. M. Pontoglio, et al. *Cell* 84, 575 (1996).
- 25 11. F.M. Sladek, Seidel, S. D. in. *Nuclear Receptors and Genetic Disease*. T. P. Burris, Ed. (Academic Press, New York, 2001).
12. F. Parviz, et al. *Nat Genet* 34, 292 (2003).
13. R.H. Costa, et al, *Hepatology* 38, 1331 (2003).
14. D.Q. Shih, et al. *Diabetes* 50, 2472 (2001).
- 30 15. D.Q. Shih, M. Stoffel. *Proc Natl Acad Sci U S A* 98, 14189 (2001).
16. K.S. Zaret. *Nat Rev Genet* 3, 499 (2002).
17. P. Jacquemin et al. *Dev Biol* 258, 105 (2003).

18. Fajans, S. S., et al. *N Engl J Med* 345, 971 (2001).
19. See supporting data on Science Online, and additional information is available at the authors' website: <http://web.wi.mit.edu/young/pancregulators>
20. H.H. Ng, F. Robert, R.A. Young, K. Struhl. *Genes Dev* 16, 806 (2002).
- 5 21. R. Bort, K. Zaret. *Nat Genet* 32, 85 (2002).
22. R. Milo, et al. *Science* 298, 824 (2002).
23. S.F. Boj et al. *Proc Natl Acad Sci U S A* 98, 14481 (2001).
24. H. Thomas, et al. *Hum Mol Genet* 10, 2089 (2001).
25. J. Ferrer. *Diabetes* 51, 2355 (2002).
- 10 26. I. Barroso et al. *PLoS Biology* 1, 41 (2003).
27. Q. Zhu et al. *Diabetologia* 46, 567 (2003).
28. L. Love-Gregory et al. *Diabetes* 54 (2004) in press.
29. K. Silander et al. *Diabetes* 54 (2004) in press.

Claims:

1. A method of determining which genes from a subset of genes are regulated by a transcriptional regulator expressed in a cell, the method comprising
- 5 (a) selectively isolating chromatin from the cell to generate isolated chromatin;
 - (b) selectively isolating chromatin fragments from the isolated chromatin to generate bound chromatin fragments, wherein the bound chromatin fragments are bound by the transcriptional regulator;
 - 10 (c) amplifying both the bound chromatin fragments to generate amplified chromatin fragments and the isolated chromatin to generate amplified control chromatin;
 - (d) hybridizing the amplified control chromatin and the amplified chromatin fragments to a DNA microarray, wherein the DNA microarray comprises
 - 15 (1) at least 10,000 experimental spots, each experimental spot comprising an experimental DNA, each experimental DNA comprising a promoter region from a gene in the subset; and
 - (2) at least 100 control spots, each control spot comprising a control DNA, each control DNA comprising a non-promoter region; and
 - 20 (e) determining and comparing a hybridization signal at each of the spots on the microarray between those generated by
 - (1) the amplified control chromatin; and
 - (2) the amplified chromatin fragments;
 - 25 wherein a gene in the subset is said to be regulated by the transcriptional regulator in the cell if a spot comprising a promoter region of said gene displays a higher level of hybridization by the amplified chromatin fragments than by the amplified control chromatin.
- 30 2. The method of claim 1, wherein the level of hybridization of the amplified chromatin fragments to each experimental spot is normalized by the level of hybridization of the amplified chromatin fragments to the control spots.

3. The method of claim 1, wherein the level of hybridization of the amplified chromatin fragments to each experimental spot is normalized by subtracting the mean level of hybridization of the amplified chromatin fragments to the control spots.
4. The method of claim 1, wherein the higher level of hybridization comprises at least a two-fold higher level of hybridization.
5. The method of claim 1, wherein the transcriptional regulator is native to the cell.
6. The method of claim 1, wherein the transcriptional regulator is not a recombinant transcriptional regulator.
7. The method of claim 1, wherein the cell is a primary cell.
8. The method of claim 7, wherein the cell is a human cell.
9. The method of claim 8, wherein the cell is a transplant-grade human cell.
10. The method of claim 1, wherein step (b) comprises immunoprecipitation of the transcriptional regulator.
11. The method of claim 1, wherein step (c) comprises ligation-mediated polymerase chain reaction (LM-PCR).
12. The method of claim 1, wherein the promoter region of the gene comprises from at least 700bp upstream to at least 200 bp downstream of the transcriptional start site of the gene.

13. The method of claim 1, wherein the promoter region comprises at least 30, 40, 50, or 60 or nucleotides in length.
14. The method of claim 1, wherein the promoter region of the gene comprises a
5 sequence of at least 30 nucleotides whose sequence is identical to a region stretching from 3 kb upstream to 1 kb downstream of the transcriptional start site of said gene.
15. The method of claim 1, wherein the non-promoter region comprises an open
10 reading frame.
16. The method of claim 1, wherein the transcriptional regulator is a basal transcription factor.
- 15 17. The method of claim 16, wherein the transcriptional regulator is an RNA polymerase II or a TATA-binding protein.
18. A method of identifying a transcriptional regulatory network in a cell, the
20 method comprising determining if a transcriptional regulator regulates additional transcriptional regulators in the cell using the method of claim 1, wherein a transcriptional regulatory network is identified if at least one additional transcriptional regulator is determined to be regulated by the transcriptional regulator.
- 25 19. The method of claim 18, wherein the experimental DNA comprises promoter regions from the additional transcriptional regulators.
20. A method of identifying a transcriptional regulatory network in a cell, the
30 method comprising determining if a transcriptional regulator regulates
(i) its own promoter; or
(ii) a promoter from a plurality of transcriptional regulators,
using the method of claim 1, wherein the experimental DNA comprises

- (a) a promoter from the transcriptional regulator; and
(b) promoters from the plurality of transcriptional regulators;
wherein a transcriptional regulatory network is identified if the transcriptional
regulator regulates itself or if it regulates at least one of the plurality of
transcriptional regulators.
21. A method of identifying transcriptional regulatory networks in a cell, the
method comprising
- (a) determining, by repeating the method of claim 1 for each of a plurality of
transcriptional regulators, the genes in a subset which are regulated by
each of the plurality of transcriptional regulators, wherein the
experimental DNA comprises promoter regions for each of the plurality
of transcriptional regulators;
- (b) determining if any one of the plurality of transcriptional regulators are
regulated by at least one of the plurality of transcriptional regulators;
wherein a transcriptional regulatory network is identified if any one of the
plurality of transcriptional regulators is regulated by at least one of the plurality
of transcriptional regulators.
22. The method of claim 21, further comprising determining if a gene is regulated
by more than one of the plurality of transcriptional regulators.
23. A DNA microarray for determining promoter occupancy in a human cell, the
microarray comprising
- (1) at least 10,000 experimental spots, each experimental spot comprising an
experimental DNA, each experimental DNA comprising a promoter
region from a human gene in the subset; and
- (2) at least 100 control spots, each control spot comprising a control DNA,
each control DNA comprising a non-promoter region;
- wherein at least 75% of the promoter regions comprise from at least 700bp
upstream to at least 200 bp downstream of the transcriptional start site.

24. A method of estimating if a transcriptional regulator is a global transcriptional regulator, the method comprising
- (a) selectively isolating chromatin from a tissue;
 - (b) identifying promoter regions from the chromatin which are bound by a candidate global transcriptional regulator;
 - (c) identifying promoter regions from the chromatin which are bound by a member of the basal transcriptional machinery; and
 - (d) comparing the promoter regions identified in steps (b) and (c) to determine the ratio between (i) the number of promoter regions bound by both the candidate global transcriptional regulator and the member of the basal transcriptional machinery; and (ii) the number of promoter regions bound by the member of the basal transcriptional machinery
- wherein a transcriptional regulator is a global transcriptional regulator when the ratio is greater than 0.2.
25. The method of claim 24, wherein steps (b) and (c) are performed using a DNA microarray.
26. The method of claim 25, wherein the DNA microarray comprises
- (i) at least 10,000 experimental spots, each experimental spot comprising an experimental DNA, each experimental DNA comprising a promoter region from a human gene in the subset; and
 - (ii) at least 100 control spots, each control spot comprising a control DNA, each control DNA comprising a non-promoter region;
27. The method of claim 24, wherein the member of the basal transcriptional machinery is an RNA polymerase II or a TATA-binding protein.
28. The method of claim 24, wherein the tissue is transplant-grade tissue.
29. The method of claim 24, wherein the tissue is freshly-isolated human tissue.

30. The method of claim 29, wherein the tissue is from a subject afflicted with a disorder.
- 5 31. The method of claim 30 wherein the disorder is a hyperplastic condition.
32. A method of identifying at least one target gene for the development of a therapeutic to treat or prevent a disorder in a subject, wherein at least one form of the disorder is caused by an altered activity in a transcriptional regulator or in
10 a suspected transcriptional regulator, the method comprising
(a) identifying the genes regulated by the transcriptional regulator in a cell;
(b) determining if the transcriptional regulator is a broad-acting transcriptional regulator or a narrow-acting transcriptional regulator, wherein if the transcriptional regulator is a broad acting transcriptional
15 regulator then the transcriptional regulator is a target gene for the development of a therapeutic, and wherein if the transcriptional regulator is a narrow acting transcriptional regulator then
(i) determining if at least one gene regulated by the transcriptional regulator is likely causative in the disorder, wherein a gene that
20 is likely causative in the disorder is a target gene for the development of a therapeutic; and
(ii) reiterating steps (a) and (b) for at least one gene that is regulated by the transcriptional regulator in the cell and that either
25 (1) encodes a transcriptional regulator or
(2) is suspected to encode a transcriptional regulator,
with the modification that the transcriptional regulator of steps (a) and (b) is said gene,
thereby identifying at least one target gene for the development of a therapeutic
30 to treat or prevent a disorder in the subject.
33. The method of claim 32, wherein identifying the genes regulated by the

transcriptional regulator in a cell comprises chromosome-wide location analysis.

34. The method of claim 32, wherein identifying the genes regulated by the transcriptional regulator in the cell comprises using the method of claim 1.
35. The method of claim 32, wherein the transcriptional regulator is a master regulatory gene.
36. The method of claim 35, wherein the master regulatory gene is SOX1-18, OCT6, PAX3, Myocardin, GATA1-6, TCF1/HNF1A, HNF4A, HNF6, NGN3, C/EBP, FOXA1-3, IPF1, GATA, HNF3, NKX2.1, CDX, FTF/NR5A2, C/EBPbeta, SCL1, SKIN1, or a member of the neurogenin, LK, LMO, SOX, OCT, PAX, GATA or MyoD family of transcription factors.
37. The method of claim 32, wherein the transcriptional regulator is PAX3, EGR-1, EGR-2, OCT6, a SOX family member, a GATA family member, a PAX family member, an OCT family member, RFX5, WHN, GATA1, VDR, CRX, CBP, MeCP2, AML1, p53, PLZF, PML, Rb, WT1, NR3C2, GCCR, PPARGgamma, SIM1, HNF1alpha, HNF1beta, HNF4alpha, PDX1, MAFA, FOXA2, or NEUROD1.
38. The method of claim 32, wherein the cell is derived from a tissue whose function is impaired in the disorder.
39. The method of the claim 32, wherein the broad acting gene regulates at least about 2.5% of the genes in the cell, and wherein the narrow acting gene regulates less than about 2.5% of the genes in the cell.
40. The method of claim 32, wherein the gene is suspected to encode a transcriptional regulator if it shares at least 30% amino acid sequence identity with the DNA binding domain of a transcriptional regulator.

41. The method of claim 32, wherein the transcriptional regulator in the cell is a mutant transcriptional regulator.
- 5 42. The method of claim 32, wherein the transcriptional regulator in the cell has altered activity.
43. The method of claim 32, wherein the gene regulated by the transcriptional regulator is likely causative of the disorder when a mutation in the gene results
10 in at least one phenotype or symptom associated with the disorder.
44. The method of claim 32, wherein the gene regulated by the transcriptional regulator is likely causative of the disorder when the gene encodes an enzyme or signaling molecule which functions in a pathway that is impaired in the
15 disorder.
45. The method of claim 32, wherein the altered activity in the transcriptional regulator comprises at least one of the following:
- 20 (a) an alteration in the binding affinity of the transcriptional regulator to DNA;
- (b) an alteration in the ability of the transcriptional regulator to bind to RNA polymerase, to an RNA polymerase holoenzyme, or to a second transcriptional regulator;
- 25 (c) an alteration in the binding affinity of the transcriptional regulator to a ligand;
- (d) an alteration in expression level or expression pattern of the transcriptional regulator; or
- (e) an alteration in an ability of the transcriptional regulator to form homomultimers or heteromultimers.
- 30 46. The method of claim 32, wherein the disorder is characterized by impaired function of at least one of the following: brain, spinal cord, heart, arteries,

esophagus, stomach, small intestine, large intestine, liver, pancreas, lungs, kidney, urinary tract, ovaries, breasts, uterus, testis, penis, colon, prostate, bone, muscle, cartilage, thyroid gland, adrenal gland, pituitary, bone marrow, blood, thymus, spleen, lymph nodes, skin, eye, ear, nose, teeth or tongue.

5

47. The method of claim 32, wherein the therapeutic comprises a small molecule drug, an antisense reagent, an antibody, a peptide, a ligand, a fatty acid, a hormone or a metabolite.

10 48. The method of claim 32, wherein the subject is a mammal.

49. The method of claim 48, wherein the mammal is a human.

50. The method of claim 32, wherein the transcriptional regulator is a
15 transcriptional activator or a transcriptional repressor.

51. The method of claim 32, wherein the transcriptional regulator is native to the cell.

20 52. The method of claim 32, wherein the transcriptional regulator is from a species different from that of the cell.

53. The method of claim 52, wherein the transcriptional regulator is a viral transcriptional regulator.

25

54. A method of treating or preventing type II diabetes in a subject, comprising administering to the subject a therapeutically effective amount of an agent that increases the global transcriptional activity of HNF4alpha.

30 55. A method of treating or preventing a disorder associated with low transcriptional activity of HNF4alpha in a subject, comprising administering to the subject a therapeutically effective amount of an agent that increases the

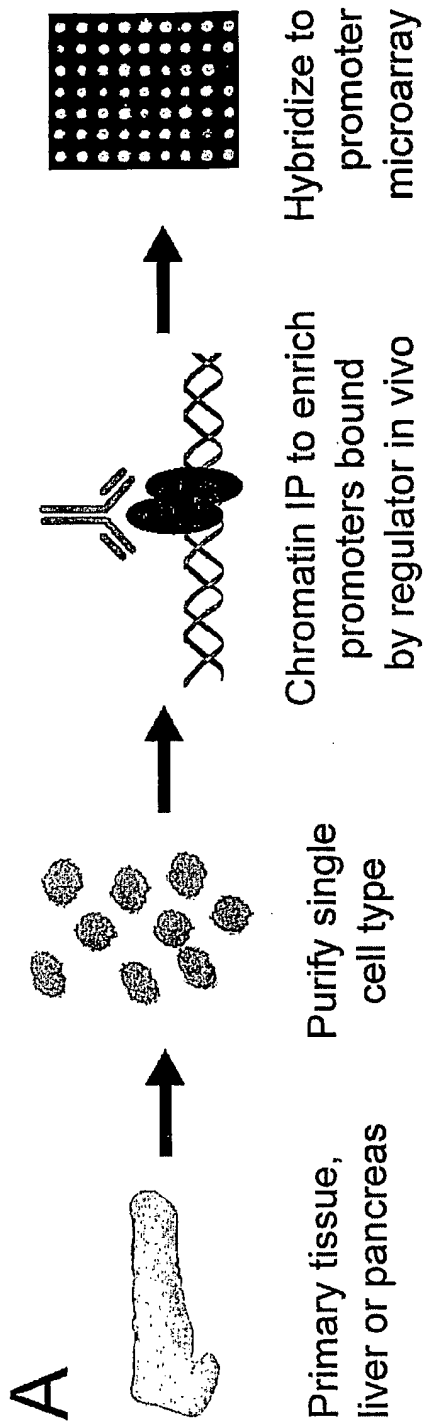
global transcriptional activity of HNF4alpha.

56. A method of treating or preventing a disorder associated with high transcriptional activity of HNF4alpha in a subject, comprising administering to the subject a therapeutically effective amount of an agent that decreases the global transcriptional activity of HNF4alpha.
57. A method of increasing the global transcriptional activity in a liver or a pancreatic cell comprising contacting the cell with an agent which increases the global transcriptional activity of HNF4alpha.
58. A method of decreasing the global transcriptional activity in a liver or a pancreatic cell comprising contacting the cell with an agent which decreases the global transcriptional activity of HNF4alpha.
59. A method of regulating the expression level of any one of the genes in Figure 13 in a hepatocyte, the method comprising contacting the cell with an agent which regulates the transcriptional activity of HNF1alpha.
60. A method of regulating the expression level of any one of the genes in Figure 14 in a pancreatic cell, the method comprising contacting the cell with an agent which regulates the transcriptional activity of HNF1alpha.
61. A method of regulating the expression level of any one of the genes in Figure 16 in a hepatocyte, the method comprising contacting the cell with an agent which regulates the transcriptional activity of HNF6.
62. A method of regulating the expression level of any one of the genes in Figure 17 in a pancreatic cell, the method comprising contacting the cell with an agent which regulates the transcriptional activity of HNF6.
63. A method of regulating the expression level of any one of the genes in Figure

18 in a hepatocyte, the method comprising contacting the cell with an agent which regulates the transcriptional activity of HNF4alpha.

64. A method of regulating the expression level of any one of the genes in Figure 19 in a pancreatic cell, the method comprising contacting the cell with an agent which regulated the transcriptional activity of HNF4alpha.
65. A method of identifying transcriptionally active genes that are regulated by a transcriptional regulator in a cell, the method comprising
- (a) selectively isolating chromatin from a tissue;
 - (b) identifying promoter regions from the chromatin that are bound by the transcriptional regulator;
 - (c) identifying promoter regions from the chromatin that are bound by a member of the basal transcriptional machinery; and
 - (d) comparing the promoter regions identified in steps (b) and (c) to determine overlapping genes,
- wherein the overlapping genes are transcriptionally active genes regulated by the transcriptional regulator.

Fig. 1A



A

Fig. 1B

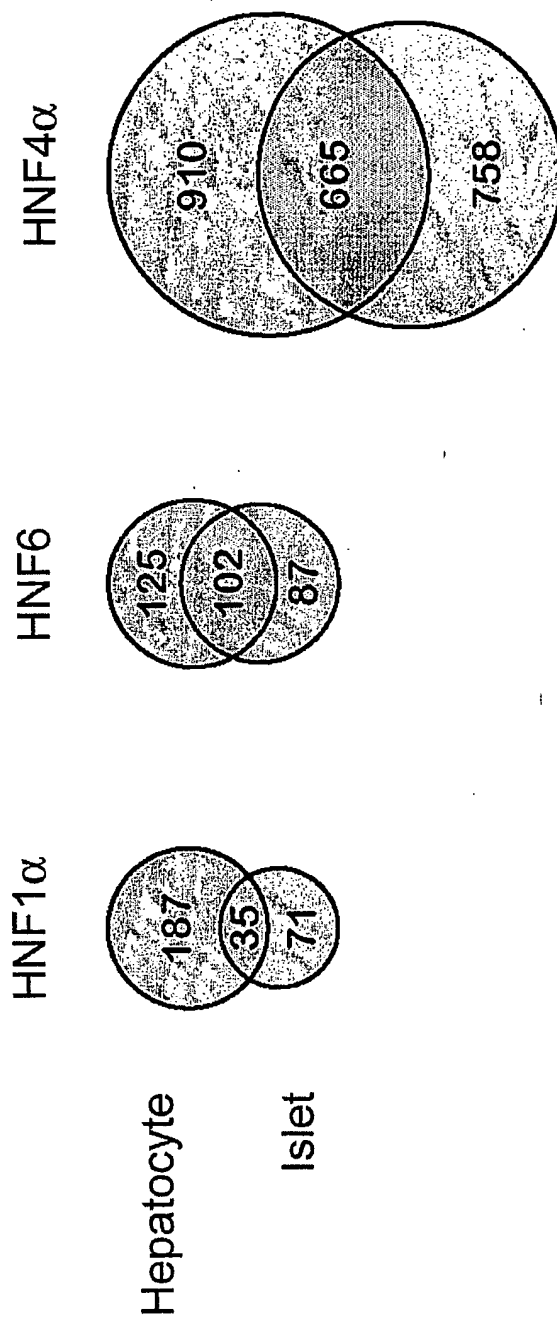
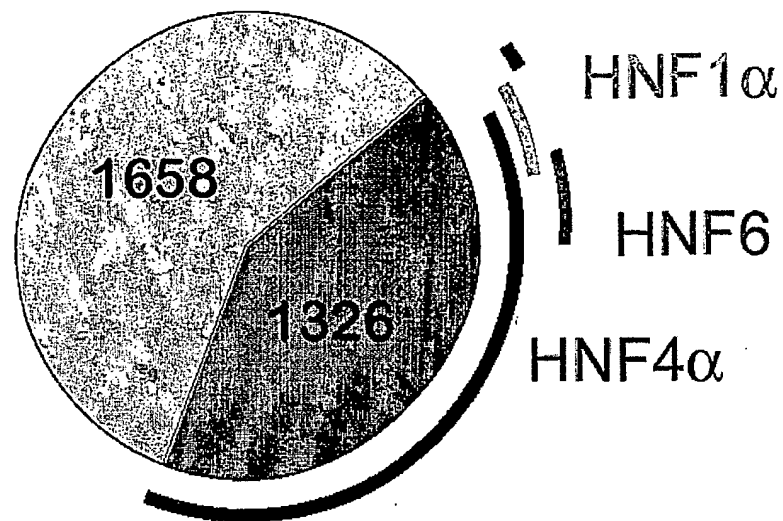
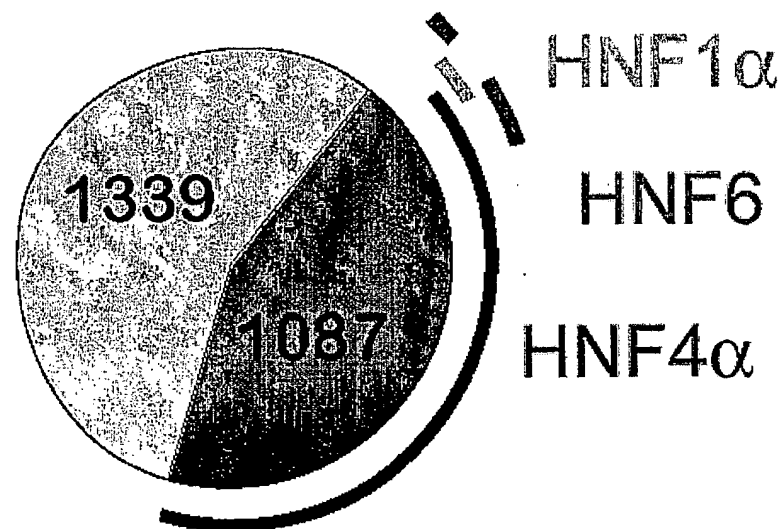


Fig. 1C



Hepatocyte



Pancreatic Islet

Fig. 2A

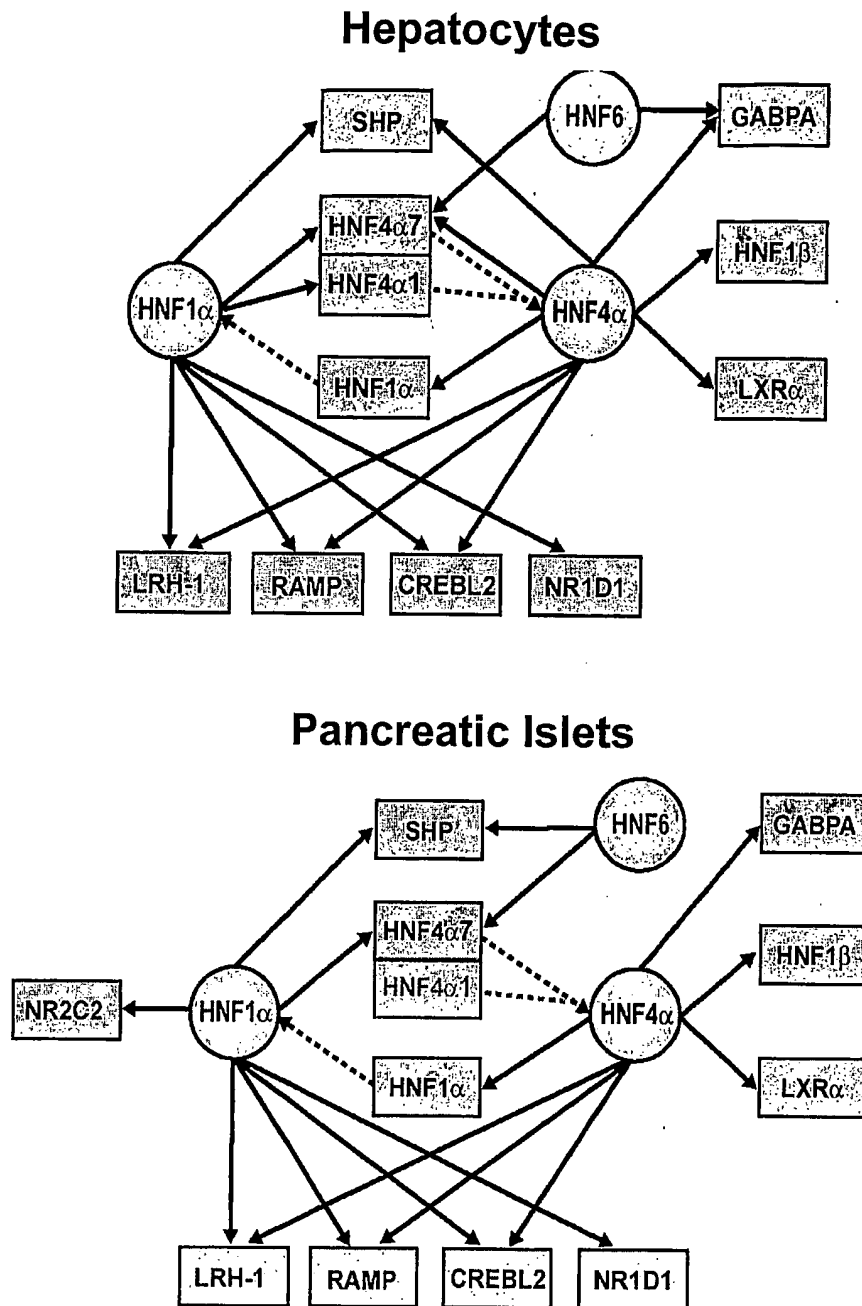


Fig. 2B

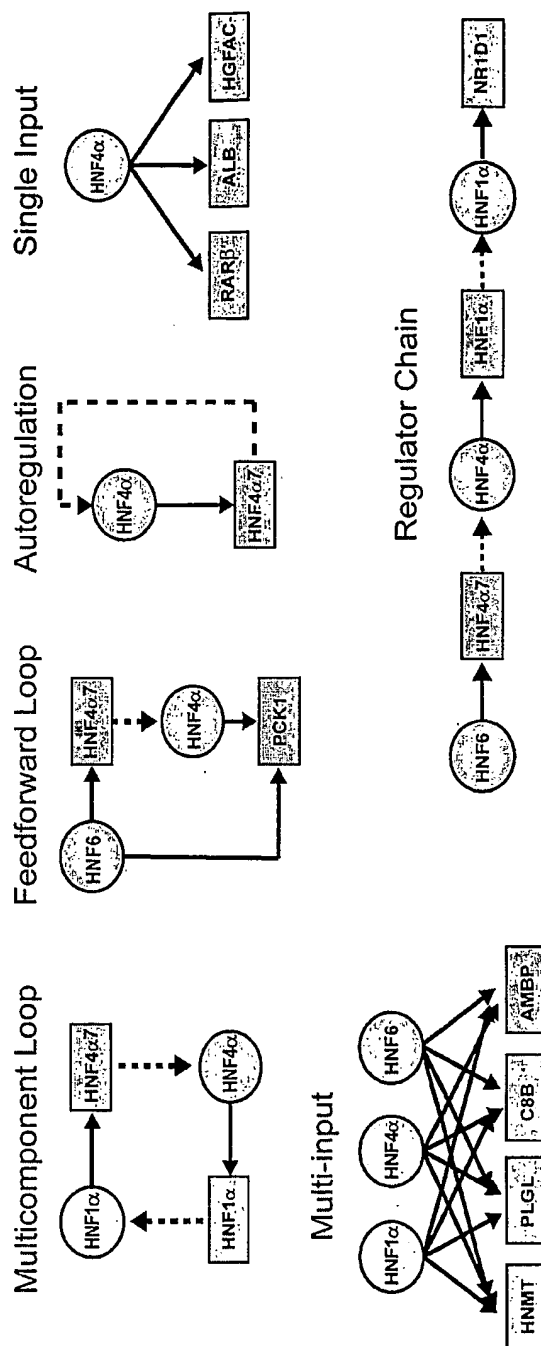


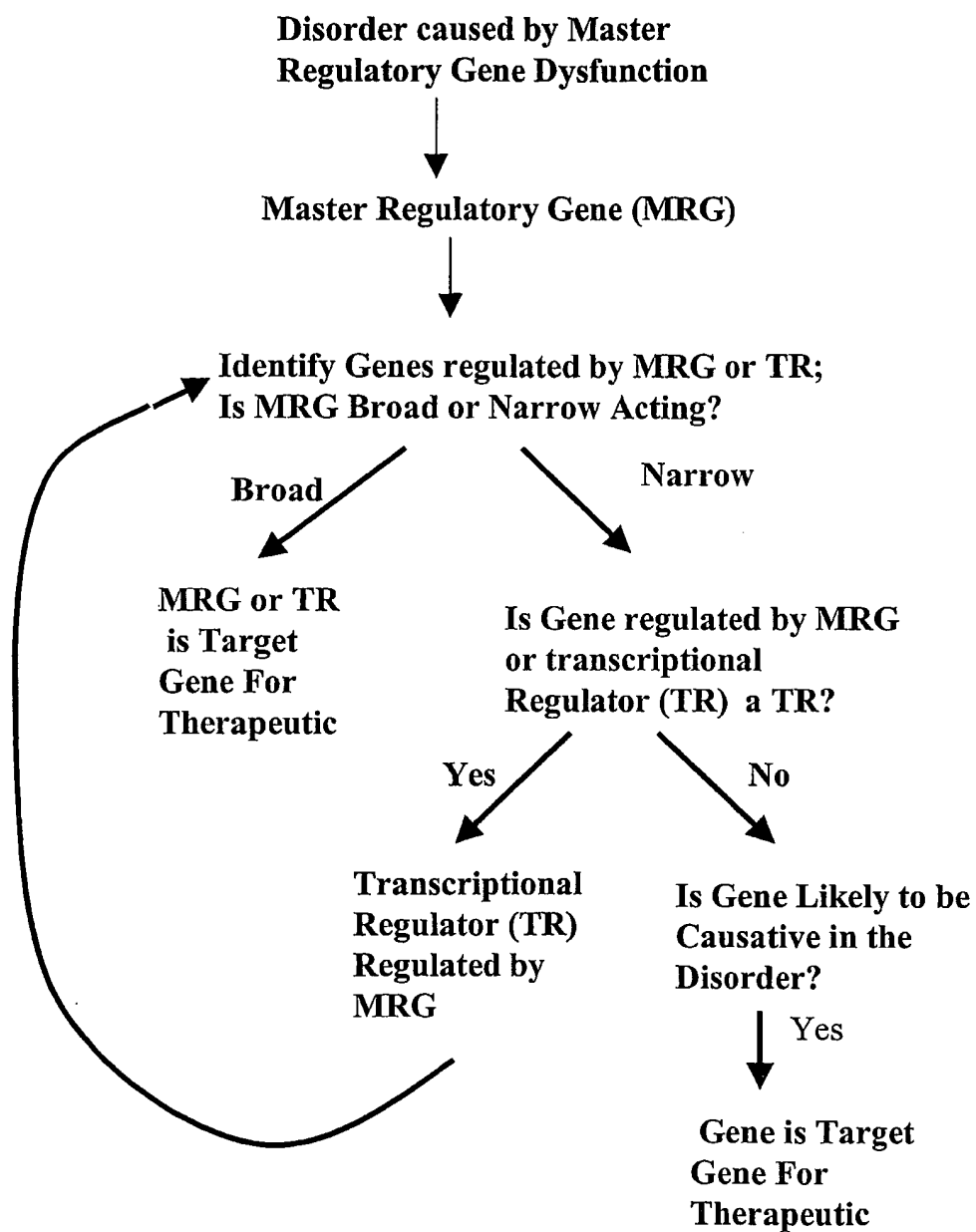
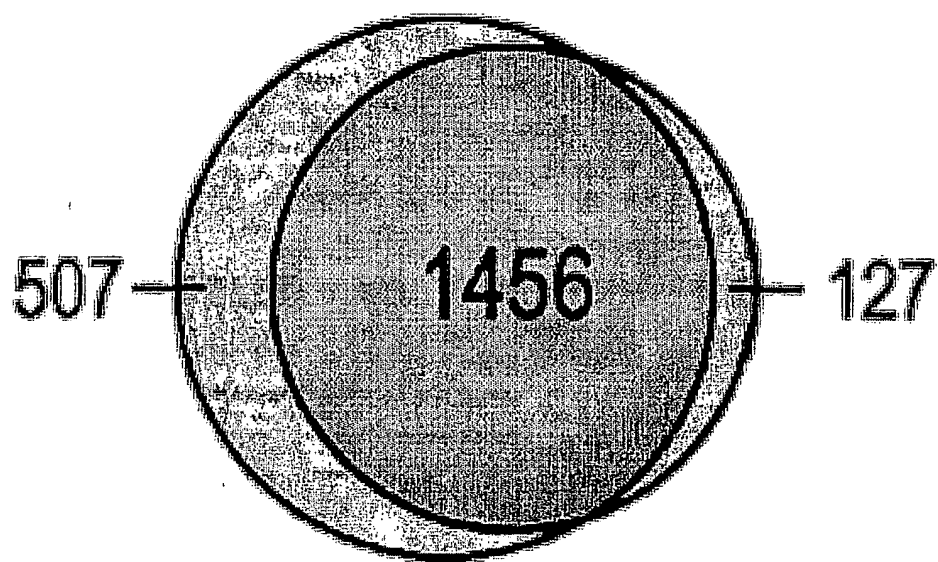

Fig. 3

Fig. 4



 sc-8987


 sc-6556

Fig. 5

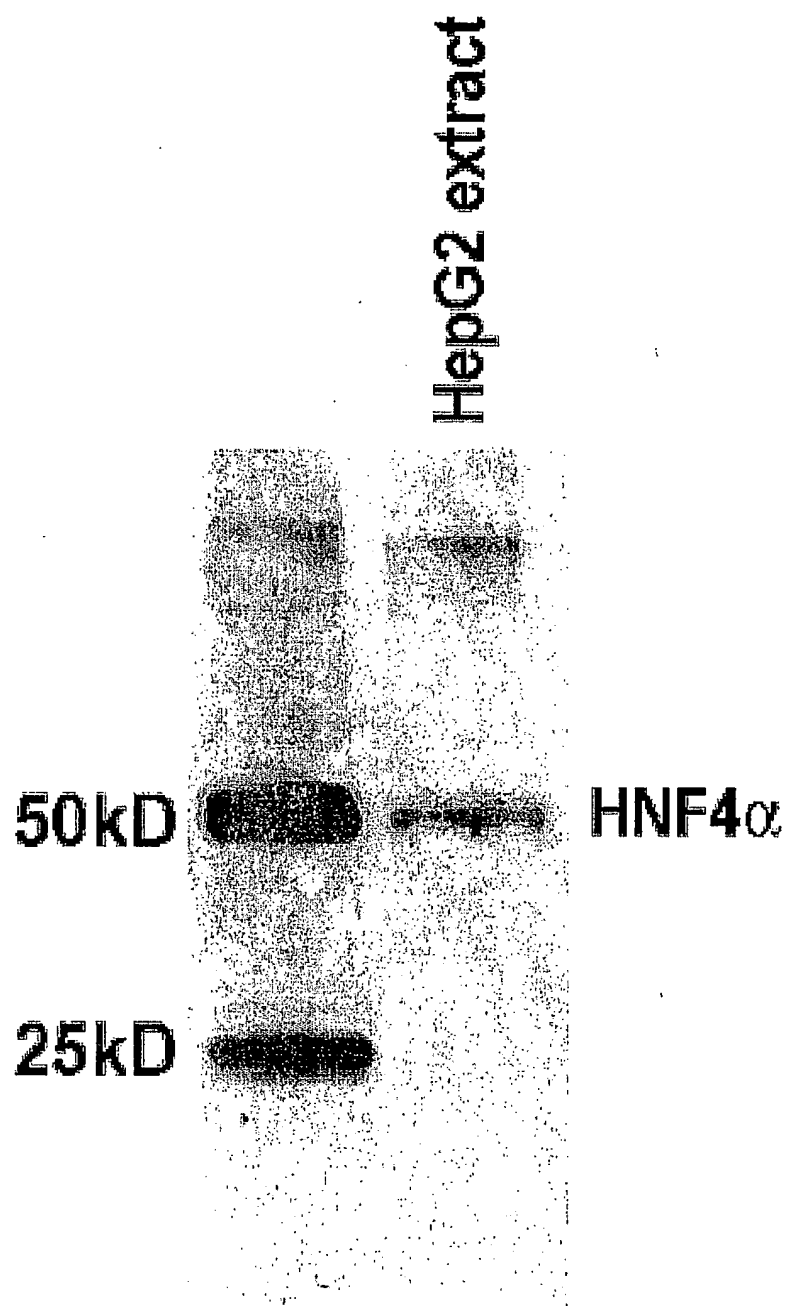


Fig. 6A

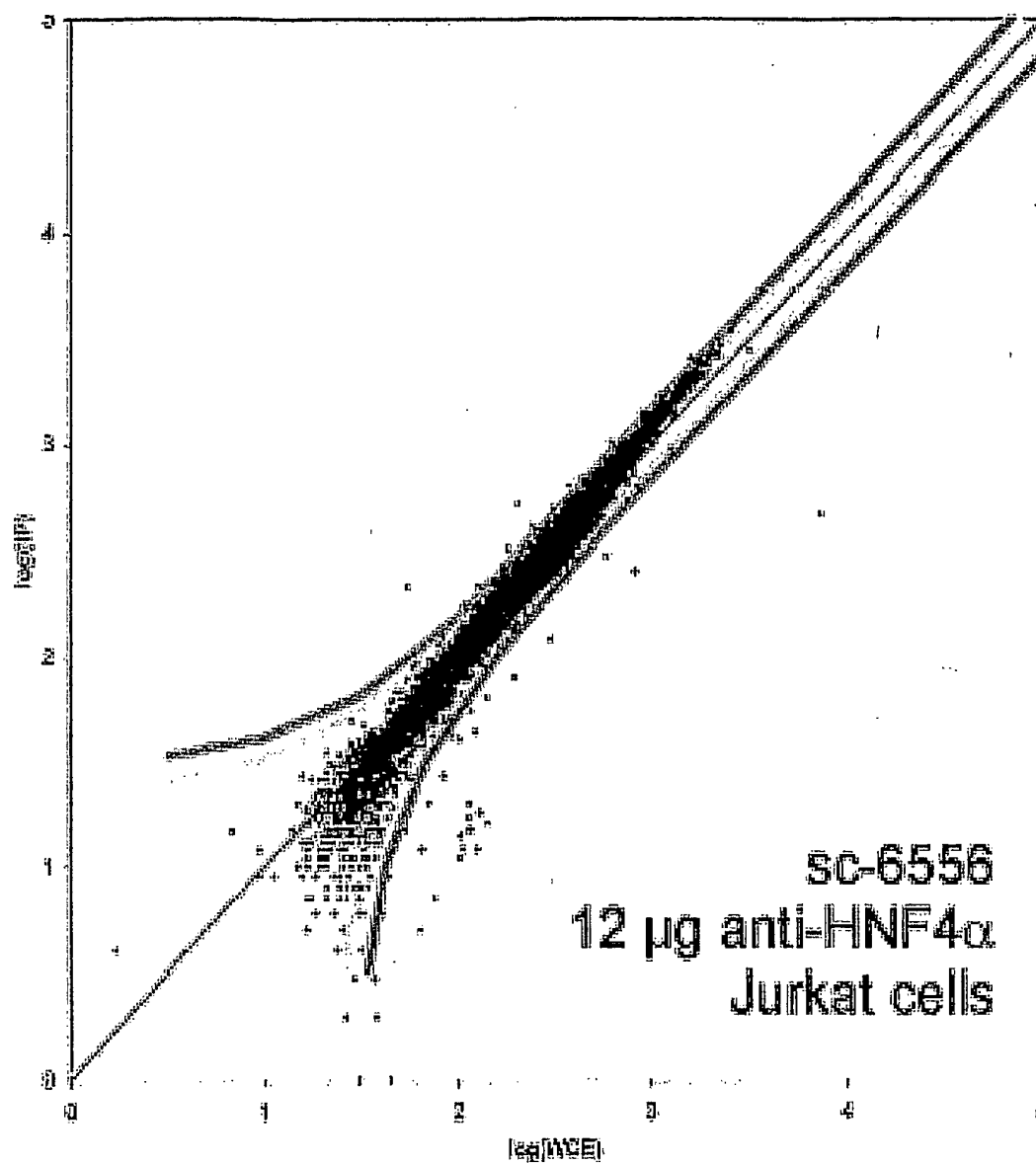


Fig. 6B

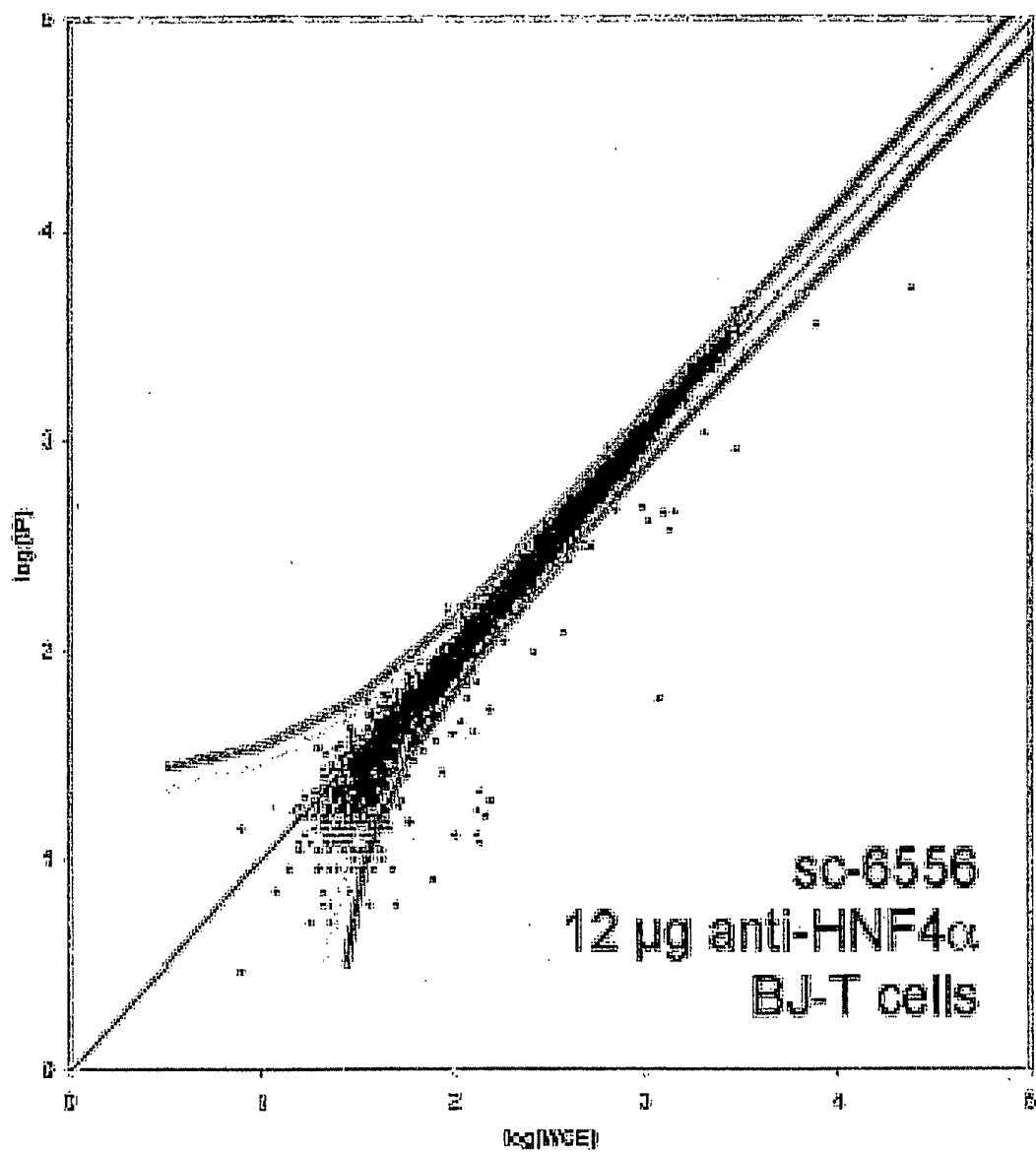


Fig. 6C

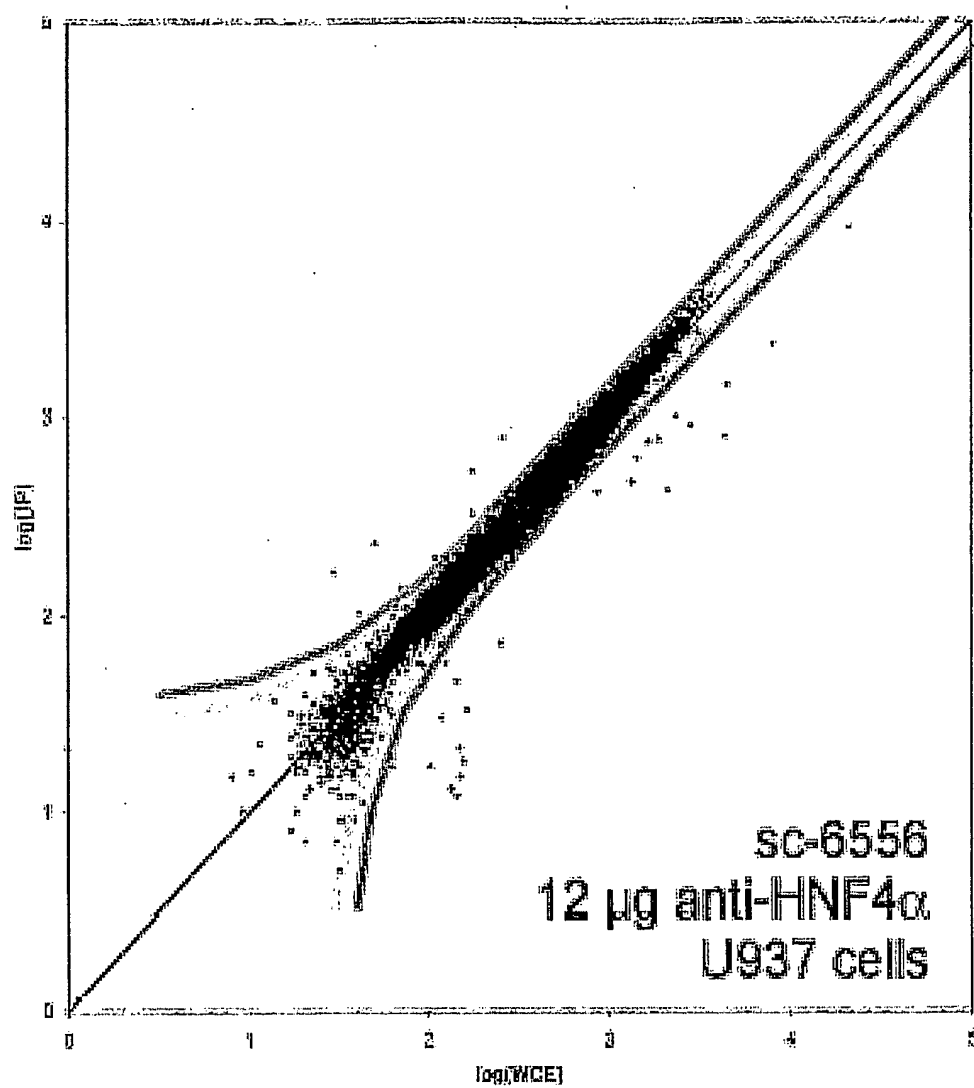


Fig. 6D

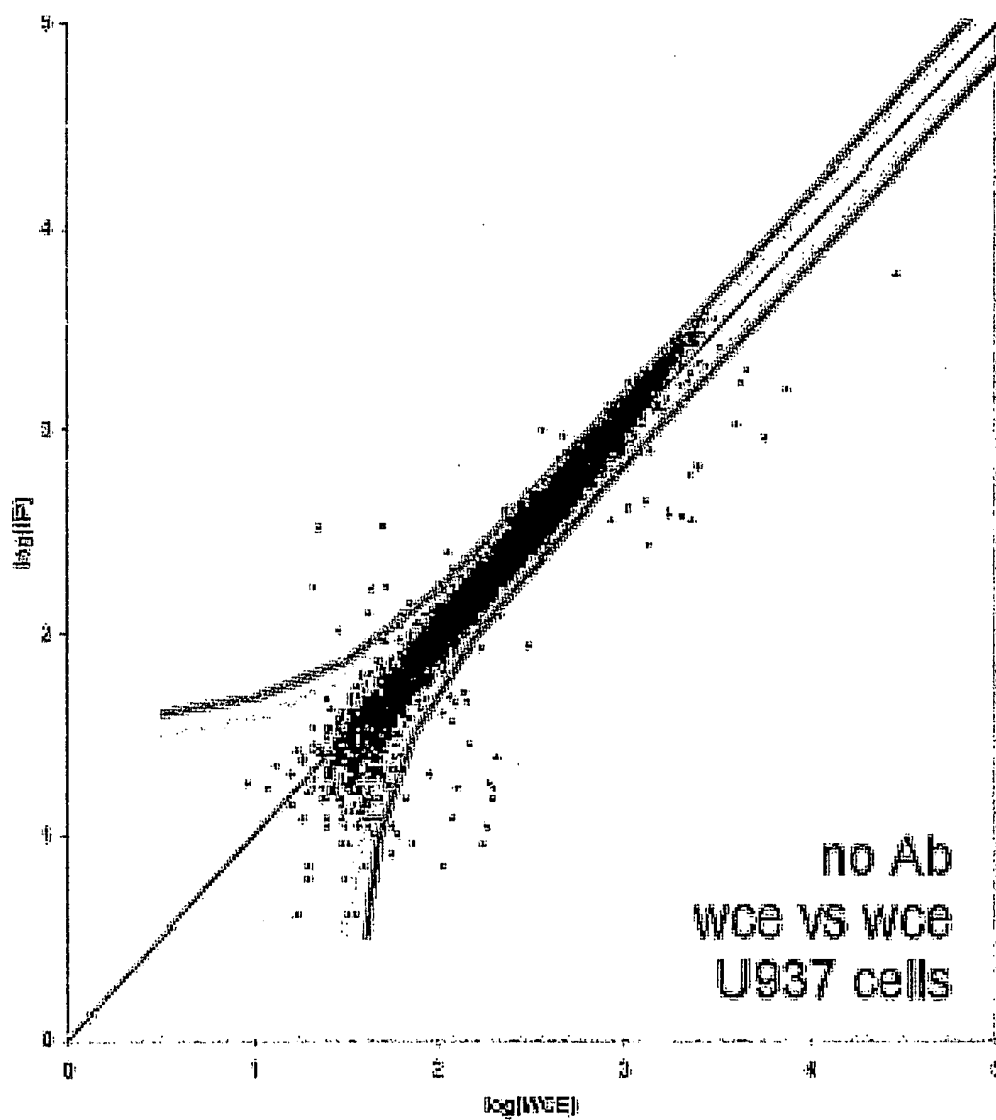


Fig. 7

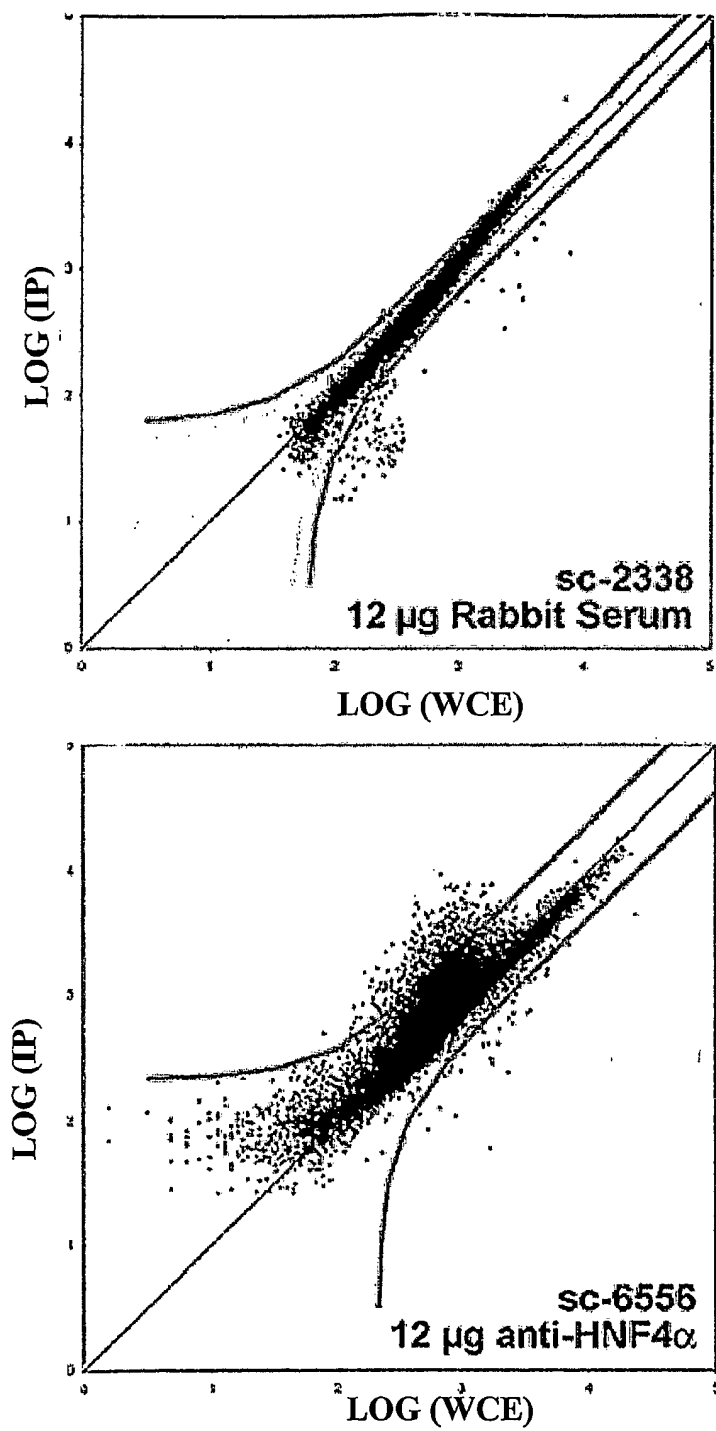


Fig. 8

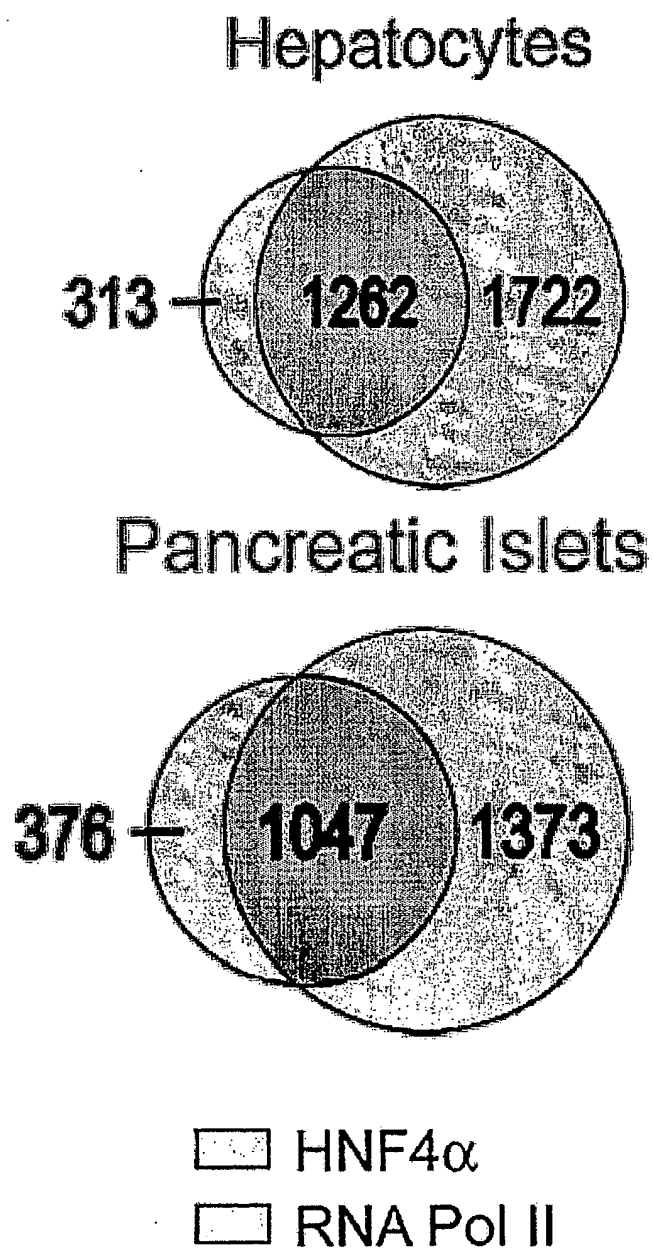


Fig. 9

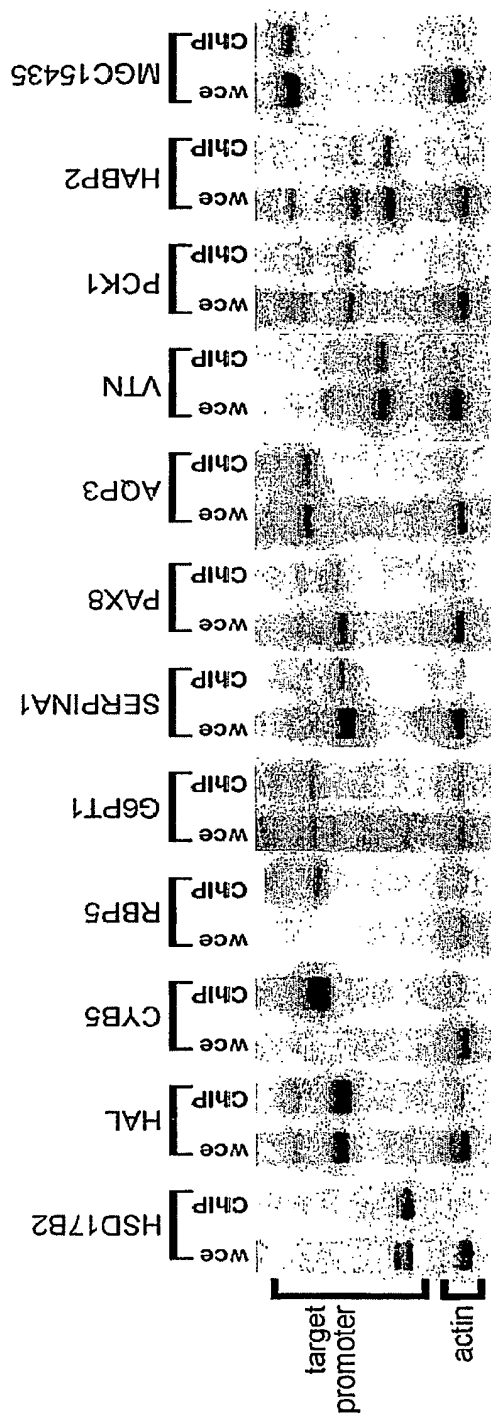


Fig. 10

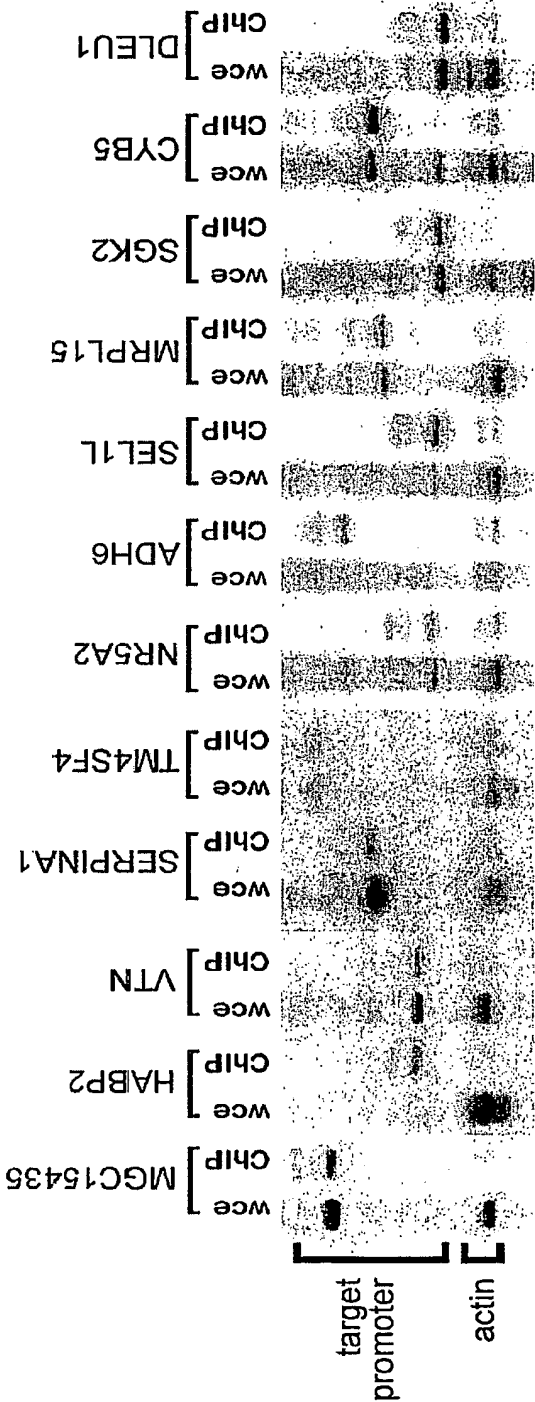


Fig. 11

Name	RefSeq	Description	Hepatocyte	Islets	Name	RefSeq	Description	Hepatocyte	Islets
Chaperone					Signal Transduction--Other				
C4BPA	NM_000715	complement 4 binding protein a	✓	✓	BIKE	NM_017593	BMP-2 inducible kinase	✓	✓
APCS	NM_001639	amyloid P component	✓		SGK2	NM_016276	serum/glucocorticoid reg. kinase 2	✓	✓
F11	NM_019559	coagulation factor XI	✓		SEL1L	NM_005065	suppressor of lin-12-like	✓	✓
C1S	NM_001734	complement component 1s	✓		SCYE1	NM_004757	small cytokine E1	✓	✓
VTN	NM_000638	somatomedin B	✓		ANGPTL3	NM_014495	angiopoietin-like 3	✓	
Enzyme--Hydrolase					Signal Transduction--Receptor				
PGCP	NM_016134	glutamate carboxypeptidase		✓	HAVCR-1	NM_012206	hepatitis A virus cellular receptor 1		✓
GLA	NM_000169	galactosidase, alpha		✓	TACR3	NM_001059	tachykinin receptor 3		✓
LIPA	NM_000235	lipase A		✓	GNB2L1	NM_006098	GTP binding protein , beta2L1		✓
SPO11	NM_012444	SPO11-like		✓	INSR	NM_000208	insulin receptor		✓
PAFAH2	NM_000437	platelet-activating factor 2	✓	✓	SSTR1	NM_001049	somatostatin receptor 1	✓	✓
AADAC	NM_001086	arylacetamide deacetylase		✓	TM4SF4	NM_004617	transmembrane 4-4	✓	✓
PS-PLA1	NM_015900	phospholipase A1alpha	✓	✓	ASGR2	NM_001181	asialoglycoprotein receptor 2	✓	✓
VNN3	NM_018399	vanin 3	✓	✓	GPR39	NM_001508	G protein-coupled receptor 39	✓	✓
CPB2	NM_016413	carboxypeptidase B2	✓		IFNAR1	NM_000629	interferon receptor 1	✓	✓
ANPEP	NM_001150	alanine aminopeptidase	✓		TFRC	NM_003234	transferrin receptor	✓	
HGFAC	NM_001528	HGF activator	✓		Transcription Regulation				
ENPEP	NM_001977	glutamyl aminopeptidase	✓		ZNF300	NM_052860	kruppel-like zinc finger protein		✓
Enzyme--Ligase					BCL6	NM_001706	B-cell CLL/lymphoma 6		✓
MCC1	NM_020166	methylcrotonoyl-CoA carboxylase		✓	ZNF155	NM_003445	zinc finger protein 155		✓
GARS	NM_002047	glycyl-tRNA synthetase	✓		FBXO8	NM_012180	F-box only protein 8		✓
TARS	NM_003191	threonyl-tRNA synthetase	✓		NR0B2	NM_021969	Small heterodimer protein	✓	✓
Enzyme--Lyase					HNF4a7	AF509467	HNF4alpha, alternate splice	✓	✓
UROD	NM_000374	uroporphyrinogen decarboxylase		✓	NR5A2	NM_003822	LHR-1/FTZ-F1	✓	✓
PCK1	NM_002591	PEPCK1	✓		ELF3	NM_004433	E74-like factor 3	✓	✓
HPCL2	NM_012260	2-hydroxyphytanoyl-CoA lyase	✓		NR1D1	NM_021724	THRA1	✓	✓
HAL	NM_002108	histidine ammonia-lyase	✓		ATF2	NM_001880	activating transcription factor 2	✓	✓
FH	NM_000143	fumarate hydratase	✓		CREBL2	NM_001310	CREB-like 2	✓	✓
Enzyme--Oxidoreductase					RARB	NM_016152	RAR-beta	✓	
COQ7	NM_016138	COQ7 coenzyme Q, 7		✓	Transporter--Channel/Pore				
ADH4	NM_000670	alcohol dehydrogenase 4		✓	SLC17A2	NM_005835	vesicular glutamate transporter	✓	
UQCRC2	NM_003366	ubiquinol-cyt. c reductase core prot. II	✓	✓	AQP3	NM_004925	aquaporin 3	✓	✓
CYB5-M	NM_030579	cytochrome b5	✓	✓	SLC22A11	NM_018484	hOAT4		✓
CYP2E	NM_000773	cytochrome P450, IIE	✓		GJB1	NM_000166	gap junction protein, beta 1	✓	
CYB5	NM_001914	cytochrome b-5	✓		Transporter--Lipids and Small Molecules				
HSD17B2	NM_002153	hydroxysteroid dehydrogenase 2	✓		APOH	NM_000042	apolipoprotein H	✓	✓
ADH1A	NM_000667	alcohol dehydrogenase 1A	✓		ALB	NM_000477	albumin	✓	✓
Enzyme--Transferase					ABCC2	NM_000392	canalicular OAT	✓	
GCNT3	NM_004751	glucosaminyl transferase 3		✓	G6PT1	NM_001467	glucose-6-phosphatase, transport	✓	
FNTB	NM_002028	farnesyltransferase beta	✓	✓	Transporter--Proteins				
HNMT	NM_006895	histamine N-methyltransferase	✓		RAB6KIFL	NM_005733	RAB6 interacting, kinesin-like		✓
GOT1	NM_002079	aspartate aminotransferase 1	✓		PEX13	NM_002618	peroxisome biogenesis factor 13		✓
UGT2B15	NM_001076	UDP glucosyltransferase 2B15	✓		TMP21	NM_006827	transmembrane trafficking protein	✓	✓
GBE1	NM_000158	glycogen branching enzyme	✓		RAB38	NM_031296	RAS oncogene	✓	✓
Enzyme Regulator					NAPA	NM_003827	alpha SNAP	✓	
SERPING1	NM_000062	C1-inhibitor	✓		AP3M1	NM_012095	adaptor-related prot. Complex	✓	
SERPINA1	NM_000295	alpha-1-antitrypsin	✓		SNX17	NM_014748	sorting nexin 17	✓	
ITIH4	NM_002218	inter-alpha inhibitor H4	✓						
AHSG	NM_001622	alpha-2-HS-glycoprotein	✓						
Ligand Binding									
TMOD2	NM_014548	tropomodulin 2		✓					
IGFBP1	NM_000596	IGF binding protein 1	✓						
MT1X	NM_005952	metallothionein 1X	✓						
CRP	NM_000567	C-reactive protein	✓						
APOA2	NM_001643	apolipoprotein A-II	✓						

Fig. 12

	BJ-T vs Hepatocytes*		BJ-T vs Pancreatic Islets*	
	BJ-T specific genes	Hepatocyte specific genes	BJ-T specific genes	Islet specific genes
HNF4 α /RNA Pol II	19/492 (4%)	996/2389 (42%)	29/546 (5%)	825/1898 (43%)
HNF1 α /RNA Pol II	2/492 (.4%)	123/2389 (5.1%)	4/546 (.9%)	32/1898 (1.7%)
HNF6/RNA Pol II	7/492 (1.4%)	105/2389 (4.4%)	3/546 (.5%)	68/1898 (3.6%)

Fig. 13

Name	RefSeq	Name	RefSeq	Name	RefSeq	Name	RefSeq	Name	RefSeq
AADAC	NM_001086	DLEU1	NM_005887	HPX	NM_000613	PHF2	NM_005392	ZNF288	NM_015642
ABCC2	NM_000392	DUSP6	NM_022652	HSD11B1	NM_005525	PIST	NM_020399	ZNF361	NM_018555
ACF	NM_014576	EIF4EBP2	NM_004096	HSD17B2	NM_002153	PLCB1	NM_015192		
ADH1A	NM_000667	ELF3	NM_004433	HSPC111	NM_016391	PLG	NM_000301		
ADH1B	NM_000668	ENPEP	NM_001977	HSPC129	NM_016396	PLGL	NM_002665		
ADH6	NM_000672	F11	NM_019559	IFNAR1	NM_000629	PS-PLA1	NM_015900		
AGT	NM_000029	FE65L2	NM_006051	IGF1R	NM_000875	PZP	NM_002864		
AHSG	NM_001622	FH	NM_000143	IGFBP1	NM_000596	RAB33B	NM_031296		
AK2	NM_001625	FKSG87	NM_032029	INADL	NM_005799	RAMP	NM_016448		
AKR1C2	NM_001354	FLJ10242	NM_018036	ITIH3	NM_002217	RARB	NM_016152		
AKR1C3	NM_003739	FLJ10276	NM_018045	ITIH4	NM_002218	RBP5	NM_031491		
AKR1C4	NM_001818	FLJ10525	NM_018126	ITM2B	NM_021999	RNGTT	NM_003800		
ALB	NM_000477	FLJ10583	NM_018148	KIAA0022	NM_014880	RPL37AP1	NG_000988		
ALDH3A2	NM_000382	FLJ10650	NM_018168	KIAA0669	NM_014779	SAC	NM_018417		
ALS2	NM_020919	FLJ10774	NM_024662	KIAA0844	NM_014951	SCYE1	NM_004757		
AMB	NM_001633	FLJ11000	NM_018295	KIAA0872	NM_014940	SEL1L	NM_005065		
ANGPTL3	NM_014495	FLJ11838	NM_024664	KIAA1041	NM_014947	SERPINA1	NM_000295		
ANPEP	NM_001150	FLJ12788	NM_022492	KNG	NM_000893	SERPINA10	NM_016186		
AP3M1	NM_012095	FLJ13448	NM_025147	LBP	NM_004139	SERPINA6	NM_001756		
APCS	NM_001639	FLJ13611	NM_024941	LOC51060	NM_015913	SERPINC1	NM_000488		
APG3	NM_022488	FLJ14356	NM_030824	LOC51096	NM_016001	SERPINE1	NM_000602		
APOA2	NM_001643	FLJ20080	NM_017657	LOC51326	NM_016632	SERPING1	NM_000062		
APOH	NM_000042	FLJ20718	NM_017939	LOC54518	NM_019043	SGK2	NM_016276		
AQP3	NM_004925	FLJ21272	NM_025032	LOC56902	NM_020143	SLC17A2	NM_005835		
AQP9	NM_020980	FLJ21934	NM_024743	LOC58486	NM_021211	SLC22A11	NM_018484		
ARHGAP11A	NM_014783	FLJ22551	NM_024708	LY6E	NM_002346	SLPI	NM_003064		
ASGR1	NM_001671	FLJ23259	NM_024727	M17S2	NM_031858	SNX17	NM_014748		
ASGR2	NM_001181	FNTB	NM_002028	M96	NM_007358	SRI	NM_003130		
ATF2	NM_001880	G0S2	NM_015714	MAGEA9	NM_005365	SSA2	NM_004600		
AUTL1	NM_032852	G3A	NM_019101	MGC10500	NM_031477	SSTR1	NM_001049		
BAT3	NM_004639	G6PT1	NM_001467	MGC11034	NM_031453	SSTR4	NM_001052		
BIKE	NM_017593	GARS	NM_002047	MGC11266	NM_024322	STRAIT11499	NM_021242		
BTN2A1	NM_078476	GBE1	NM_000158	MGC13010	NM_032687	SUPV3L1	NM_003171		
C1S	NM_001734	GCKR	NM_001486	MGC15435	NM_032367	SYN3	NM_133632		
C2	NM_000063	GDI2	NM_001494	MGC955	NM_024097	TARS	NM_003191		
C4BPA	NM_000715	GIOT-2	NM_016264	MIA2	NM_054024	TBPL1	NM_004865		
C8B	NM_000066	GJB1	NM_000166	MRPL15	NM_014175	TEF	NM_003216		
CCNE1	NM_001238	GOT1	NM_002079	MRPS18B	NM_014046	TFRC	NM_003234		
CDCA1	NM_031423	GPR39	NM_001508	MSH6	NM_000179	TIEG2	NM_003597		
CISH	NM_013324	GPX2	NM_002083	MT1H	NM_005951	TIEG2	NM_003597		
CLYBL	NM_138280	GRHPR	NM_012203	MT1L	NM_002450	TM4SF4	NM_004617		
CNTNAP2	NM_014141	GTF2B	NM_001514	MT1X	NM_005952	TMEM1	NM_003274		
CPB2	NM_016413	GTF2E1	NM_005513	MTHFD1	NM_005956	TNFRSF6	NM_000043		
CREBL2	NM_001310	GTPBG3	NM_032620	MTP	NM_000253	UGT1A1	NM_000463		
CRP	NM_000567	HABP2	NM_004132	NAPA	NM_003827	UGT2B11	NM_001073		
CTS2	NM_001336	HAL	NM_002108	NET-2	NM_012338	UGT2B15	NM_001076		
CYB5	NM_001914	HAO1	NM_017545	NFKB1B	NM_002503	UQCRC2	NM_003366		
CYB5-M	NM_030579	HCAP-G	NM_022346	NPC1L1	NM_013389	VNN3	NM_018399		
CYP2E	NM_000773	HGD	NM_000187	NR0B2	NM_021969	VTN	NM_000638		
CYP3A43	NM_022820	HGFAC	NM_001528	NR1D1	NM_021724	WBP4	NM_007187		
DAF	NM_000574	HNF4A	NM_000457	NR5A2	NM_003822	WDF2	NM_052950		
DC13	NM_020188	HNF4A	NM_000457	NRD1	NM_002525	WDR12	NM_018256		
DKFZP564O0463	NM_014156	HNF4a7	AF509467	PAFAH2	NM_000437	XDH	NM_000379		
DKFZP586A0522	NM_014033	HNMT	NM_006895	PAX8	NM_013952	XPC	NM_004628		
DKFZP586M0122	NM_015425	HPCL2	NM_012260	PCK1	NM_002591	ZK1	NM_005815		

Fig. 14

Name	RefSeq	Name	RefSeq
AADAC	NM_001086	KIAA0101	NM_014736
ABCC9	NM_020297	KIAA0399	NM_015113
ADH4	NM_000670	KIAA0844	NM_014951
APOH	NM_000042	KIF13A	NM_022113
ARHGAP11A	NM_014783	KIR-023GB	NM_015868
B29	NM_031939	KIR2DS2	NM_012312
BCL6	NM_001706	KIR3DL1	NM_013289
BIKE	NM_017593	KRTAP1.1	NM_030967
C4BPA	NM_000715	KRTHA3A	NM_004138
C6orf11	NM_005452	LIPA	NM_000235
CDC45L	NM_003504	LOC113201	NM_138423
COL3A1	NM_000090	LOC113220	NM_138424
COQ7	NM_016138	LOC51092	NM_015996
CPXCR1	NM_033048	LOC56906	NM_020147
CRH	NM_000756	MCCC1	NM_020166
CTS2	NM_001336	MGC10500	NM_031477
CYB5-M	NM_030579	MGC15677	NM_032878
DKFZP564J157	NM_018457	MIA2	NM_054024
DLEU1	NM_005887	MRPL15	NM_014175
DOCK1	NM_001380	Nod1(-)6kb	NM_006092
DSC1	NM_024421	NPY2R	NM_000910
EIF3S6	NM_001568	NR0B2	NM_021969
ELF3	NM_004433	NR2C2	NM_003298
FBXO8	NM_012180	NR5A2	NM_003822
FE65L2	NM_006051	PAFAH2	NM_000437
FIL1(EPSILON)	NM_014440	PAX8	NM_013952
FLJ10242	NM_018036	pcnp	NM_020357
FLJ10252	NM_018040	PEX13	NM_002618
FLJ10474	NM_018104	PGCP	NM_016134
FLJ10650	NM_018168	PRO2032	NM_018615
FLJ11301	NM_018385	PSMA5	NM_002790
FLJ13273	NM_024751	PS-PLA1	NM_015900
FLJ13385	NM_024853	RAB33B	NM_031296
FLJ13448	NM_025147	RAB6KIFL	NM_005733
FLJ14855	NM_033210	SDCCAG10	NM_005869
FLJ20156	NM_017691	SEL1L	NM_005065
FLJ20225	NM_019062	SGK2	NM_016276
FLJ20234	NM_017720	SLC26A7	NM_052832
FLJ20298	NM_017752	SPO11	NM_012444
FLJ20643	NM_017916	SRI	NM_003130
FLJ20731	NM_017946	SSTR1	NM_001049
FLJ21272	NM_025032	TACR3	NM_001059
FLJ22559	NM_024928	TM4SF4	NM_004617
FNTB	NM_002028	TMOD2	NM_014548
GCNT3	NM_004751	TMP21	NM_006827
GIOT-2	NM_016264	UQCRC2	NM_003366
GLA	NM_000169	UROD	NM_000374
GNB2L1	NM_006098	VNN3	NM_018399
GPR74	NM_004885	WBP4	NM_007187
H4F2	NM_003548	ZNF155	NM_003445
HAVCR-1	NM_012206	ZNF300	NM_052860
HHLA2	NM_007072		
HNF4a7	AF509467		
IFNA10	NM_002171		
INSR	NM_000208		

Fig. 15A

Regulator	Target Gene	Direct Reference	In vitro Reference	Indirect Reference	Sequence Based Reference	ORGANISM
HNF4 α	GST-YA			Paulson 1990		human
HNF4 α	TTR		Sladek 1990	Sladek 1990, costa 1991		human
HNF4 α	ApoC3		Sladek 1990	Sladek 1990		human
HNF4 α	ApoA1		Sladek 1990	Sladek 1990		human
HNF4 α	serpina		Sladek 1990	Sladek 1990		human
HNF4 α	Pklr		Sladek 1990	Sladek 1990		human
HNF4 α	cyp2c13				eguchi 1991	rat
HNF4 α	alb		herbst 1991	herbst 1991		rat
HNF4 α	tlr		herbst 1991	herbst 1991		rat
HNF4 α	hnf1a			tian 1991		human
HNF4 α	l9		crossley 1991			human
HNF4 α	hnf1a			kuo 1992		human
HNF4 α	apob		ladias 1992	ladias 1992		human
HNF4 α	ApoC3		ladias 1992	ladias 1992		human
HNF4 α	apoa2		ladias 1992	ladias 1992		human
HNF4 α	pklr			puzanal 1992		human
HNF4 α	l9			reljnen 1992		human
HNF4 α	lf			schaeffer 1993		human
HNF4 α	hnf1a			zapp 1993		xenopus
HNF4 α	pck1		angrand 1994	angrand 1994		rat
HNF4 α	pck2		angrand 1994	angrand 1994		rat
HNF4 α	cyp2c2		chen 1993	chen 1993		human
HNF4 α	cyp2c1		chen 1993	chen 1993		human
HNF4 α	cyp2c3		chen 1993	chen 1993		human
HNF4 α	cyp7a1		chiang 1994	chiang 1994		rat
HNF4 α	ApoA1		fuernkranz 1994	fuernkranz 1994		human
HNF4 α	CEACAM1		hauck 1994	hauck 1994		human
HNF4 α	apoa4		klistaki 1994	klistaki 1994		human
HNF4 α	pklr			liimatta 1994		rat
HNF4 α	a2m		matthijs 1994			human
HNF4 α	pklr	mlquerol 1994				human
HNF4 α	rpb2			nakshatri 1994		rodent
HNF4 α	olc			nishiyori 1994		mouse
HNF4 α	acox1		winrow 1994	winrow 1994		rat
HNF4 α	hsd17b4		winrow 1994	winrow 1994		rat
HNF4 α	l7		erdmann 1995, greenberg 1995	erdmann 1994, greenberg 1995		human
HNF4 α	l8		figueiredo 1995	figueiredo 1995		human
HNF4 α	apo		galson 1995	galson 1995		human
HNF4 α	cyp2c9		ibeanu 1995	ibeanu 1995		human
HNF4 α	ambp		rouet 1995	rouet 1995		human
HNF4 α	cyp2c23		roussel 1995			rat
HNF4 α	cyp2d6		caims 1996	caims 1996		human
HNF4 α	serpinc1		Fernandez-Rachubinski 1996	Fernandez-Rachubinski 1996		human
HNF4 α	bf			garner 1996		human
HNF4 α	fl10		hung 1996	hung 1996		human
HNF4 α	pfr		moltrup 1996	moltrup 1996		rat
HNF4 α	mst1		waltz 1996	waltz 1996		human
HNF4 α	lipc			chang 1997		human
HNF4 α	g6pc		lin 1997	lin 1997		human
HNF4 α	SLC2A2			stoffel 1997		mouse
HNF4 α	aldob			stoffel 1997		mouse
HNF4 α	gadp			stoffel 1997		mouse
HNF4 α	fabp1			stoffel 1997		mouse
HNF4 α	cyp2a4		yokomori 1997			mouse
HNF4 α	fl12		farselli 1998			human
HNF4 α	cyp3a23		huss 1998	huss 1998		rat
HNF4 α	shbg		janne 1998	janne 1998		human
HNF4 α	apoc2		kardassis 1998	kardassis 1998		human
HNF4 α	alp			magee 1998		human
HNF4 α	HMGCS2		rodriguez 1998	rodriguez 1998		rodent
HNF4 α	ALDH3A1		boesch 1999	boesch 1999		rat
HNF4 α	serpina1		hu 1999	hu 1999		human
HNF4 α	cyp3a1			ogino 1999		rat
HNF4 α	aldh2			pinalre 1999		human
HNF4 α	cyp2c12		sasaki 1999	sasaki 1999		rat
HNF4 α	GUCY2C		swenson 1999	swenson 1999		human
HNF4 α	ang		yanai 1999	yanai 1999		human
HNF4 α	ada		dusing 2000			human
HNF4 α	hnl6		lahuna 2000	lahuna 2000		human

Fig. 15B

TABLE S4		Direct	In vitro	Indirect	Sequence Based	
Regulator	Target Gene	Reference	Reference	Reference	Reference	Organism
HNF4 α	hadhb		nicolas-frances 2000	nicolas-frances 2000		human
	pax4		smith 2000	smith 2000		human
	ins			wang 2000		mouse
HNF4 α	ogdh			wang 2000		mouse
HNF4 α	ucp2			wang 2000		mouse
HNF4 α	hnf4a		bailly 2001	bailly 2001		human
HNF4 α	ghr		jiang 2001	jiang 2001		bovine
HNF4 α	cyp3a4			jover 2001		human
HNF4 α	cyp3a5			jover 2001		human
HNF4 α	cyp3a6			jover 2001		human
HNF4 α	cyp2b6			jover 2001		human
HNF4 α	cyp2c9			jover 2001		human
HNF4 α	lmo1			luo 2001		rabbit
HNF4 α	cyp3a16		nakayama2001	nakayama2001		mouse
HNF4 α	akr1c4		ozeki 2001	ozeki 2001		human
HNF4 α	cyp8b1		zhang 2001	zhang 2001		human
HNF4 α	hpd		aarenstrup 2002	aarenstrup 2002		rat
HNF4 α	cyp27		garuti 2002	garuti 2002		human
HNF4 α	NOS2A		guo 2002	guo 2002		rat
HNF4 α	cpl1a			louet 2002		human
HNF4 α	ppara		pineda-torra 2002	pineda-torra 2002		human
HNF4 α	gk		roth 2002			rat
HNF4 α	Serpina1	Soutoglou 2002				human
HNF1 α	FGA			baumhueter 1990		
HNF1 α	FGB			baumhueter 1990		
HNF1 α	FGG			baumhueter 1990		
HNF1 α	afp			baumhueter 1990		
HNF1 α	serpina1			baumhueter 1990		
HNF1 α	afm			herbst 1991	cereghini 1990	rat (herb
HNF1 α	afm			tronche 1991		rat
HNF1 α	cyp2e1		gonzalez 1990, hayashi 1991			animal
HNF1 α	aldob		raymondjean 1991			rat
HNF1 α	aldob		ito 1990			rat
HNF1 α	igfbp1			suwanichkul 1990, babajko 1993		human
HNF1 α	igfbp1			powell 1993		human
HNF1 α	igfbp1			suh 1995, suh 1997		rat
HNF1 α	crp			lonlati 1990		
HNF1 α	apoa2			chambaz 1991		human
HNF1 α	lfr			costa 1991		mouse
HNF1 α	lfr			herbst 1991		rat
HNF1 α	hdldp				drewes 1991	xenopus
HNF1 α	rbp5			tripodi 1991		human
HNF1 α	f2		bancroft 1992	bancroft 1992		human
HNF1 α	apob		brooks 1992			human
HNF1 α	lnsr		cameron 1992			human
HNF1 α	lnsr		cameron 1992			human
HNF1 α	agl			congiu 1992		mouse
HNF1 α	lms			emens 1992		rat
HNF1 α	pklr		puzenat 1992			
HNF1 α	lat		schweizer-groyer 1992			rat
HNF1 α	sial1		svensson 1992			human
HNF1 α	adh1			svensson 1992, bois-joyeux 1995		human
HNF1 α	crhbp			van ooji 1992		human
HNF1 α	afp			bernier 1993	behan 1993	human
HNF1 α	fgb		dalmon 1993	dalmon 1993		human
HNF1 α	lyz				grajer 1993	chicken
HNF1 α	aldob			gregori 1993		
HNF1 α	lbg			hayashi 1993		human
HNF1 α	apoa1			kritiis 1993		
HNF1 α	apoc3			kritiis 1993		
HNF1 α	crp		li 1996	ku 1993, li 1996		mouse
HNF1 α	fgb			roberts 1993		xenopus
HNF1 α	proc			berg 1994		human
HNF1 α	serpina1			bullia 1994		
HNF1 α	gst2		clairmont 1994			human
HNF1 α	cyp2c13			legrauerend 1994		human
HNF1 α	pklr	miqueral 1994				human
HNF1 α	anpep		olsen 1994	olsen 1994		human
HNF1 α	sl		wu 1994	wu 1994		human

Fig. 15C

TABLE S4		Direct	In vitro	Indirect	Sequence Based	
Target Gene	Reference	Reference	Reference	Reference	Reference	Organism
HNF1 α				arenzana 1995		human
HNF1 α				hu 1995		human
HNF1 α				kufik 1995, nollen 1995		salmon, human
HNF1 α			liu 1995	liu 1995, lerche 1996		rat
HNF1 α			rouet 1995	rouet 1995		human
HNF1 α			aguanno 1996	aguanno 1996		human
HNF1 α			mcglynn 1996	mcglynn 1996		human
HNF1 α			meroni 1996	meroni 1996		human
HNF1 α				pontoglio 1996		mouse
HNF1 α					boukafiane 1997	human
HNF1 α				chang 1997		rat
HNF1 α			dogra 1997	dogra 1997		chicken
HNF1 α			hansen 1997	hansen 1997		human, rat
HNF1 α			hochman 1997	hochman 1997		mouse
HNF1 α			lin 1997	lin 1997		human
HNF1 α			McGehee 1997	McGehee 1997		rodent
HNF1 α			Pontoglio 1997			mouse
HNF1 α			Taylor 1997			mouse
HNF1 α				bailly 1998		rat
HNF1 α				bailly 1998		rat
HNF1 α				bailly 1998		rat
HNF1 α			lin 1999	lin 1998		human
HNF1 α			magee 1998	magee 1998		human
HNF1 α			rhoads 1998			rat
HNF1 α				rodoloso 1998		human
HNF1 α			song y 1998	song y 1998		human
HNF1 α			song c 1998	song c 1998		rat
HNF1 α				spek 1998		human
HNF1 α			streeper 1998	streeper 1998		human
HNF1 α			trauner 1998			human
HNF1 α				wang 1998		human
HNF1 α						human, mouse
HNF1 α			bernard 1999			human
HNF1 α			chen 1999			human
HNF1 α				erickson 1999		human
HNF1 α			hu 1999	hu 1999		human
HNF1 α				meton 1999		salmon
HNF1 α			okita 1999	okita 1999		human
HNF1 α			rao 1999	rao 1999		rat
HNF1 α			spodsborg 1999	spodsborg 1999		mice
HNF1 α				wood 1999		human
HNF1 α				akiyama 2000		mouse
HNF1 α			antes 2000	antes 2000		mice
HNF1 α			cha 2000	cha 2000		human
HNF1 α			erickson 2000	erickson 2000		human
HNF1 α			gregory 2000	gregory 2000		human
HNF1 α			ishii 2000	ishii 2000		human
HNF1 α			metz 2000	metz 2000		rat
HNF1 α				muppala 2000		mouse
HNF1 α			piechocki 2000	piechocki 2000		human
HNF1 α			Pontoglio 2000	pontoglio 2000		human
HNF1 α			smith 2000	smith 2000		human
HNF1 α				wang 2000		rat
HNF1 α				wang 2000		rat
HNF1 α				wang 2000		rat
HNF1 α				wang 2000		rat
HNF1 α				wang 2000		rat
HNF1 α				wang 2000		rat
HNF1 α			bailly 2001	bailly 2001		human
HNF1 α			ben-shushan 2001	ben-shushan 2001		human
HNF1 α	Boj 2001					mouse
HNF1 α	Boj 2001					mouse
HNF1 α	Boj 2001					mouse
HNF1 α			cha 2001	cha 2001		human
HNF1 α			halzis 2001	halzis 2001		human
HNF1 α				hiraiwa 2001		mouse
HNF1 α				hiraiwa 2001		mouse
HNF1 α			jung 2001	jung 2001		human
HNF1 α				jung 2001		human
HNF1 α				lee 2001		human
HNF1 α				leu 2001		rodent
HNF1 α				leu 2001		rodent
HNF1 α				leu 2001		rodent
HNF1 α			luo 2001	luo 2001		rabbit, human

Fig. 15D

TABLE S4		Direct	In vitro	Indirect	Sequence Based
Regulator	Target Gene	Reference	Reference	Reference	Reference
HNF1 α	CYP27A1		memom 2001		hamster
HNF1 α	AKR1C4		ozeki 2001	ozeki 2001	human
HNF1 α	NR5A2		pare 2001	pare 2001	mouse
HNF1 α	cyp2c11		park 2001	park 2001	rodent
HNF1 α	cyp2a2		park 2001	park 2001	rodent
HNF1 α	cyp4a2		park 2001	park 2001	rodent
HNF1 α	pklr		parrizas 2001		human
HNF1 α	slc2a2		parrizas 2001		human
HNF1 α	pah		parrizas 2001		human
HNF1 α	c8a			pontoglio 2001	mouse
HNF1 α	c5			pontoglio 2001	mouse
HNF1 α	cyp2e1		roe 2001		rat
HNF1 α	nr1h4		shih 2001	shih 2001	mouse
HNF1 α	SLC10A2		shih 2001	shih 2001	mouse
HNF1 α	SLC17A1			soumounou 2001	human, mouse
HNF1 α	hnf4a7			thomas 2001	human
HNF1 α	ins			yamakawa 2001	human
HNF1 α	Nr5a2			zhang 2001	
HNF1 α	SLC5A1			vayro 2001	sheep
HNF1 α	slc2a2		ban 2002	ban 2002	human
HNF1 α	si			boudreau 2002	mouse
HNF1 α	SLC17A1			cheret 2002	mouse
HNF1 α	SLC10A1		geler 2002		rat
HNF1 α	UGT2B17		gregory 2002	gregory 2002	human
HNF1 α	hnf4a7			hansen 2002	mouse
HNF1 α	gjb1			koffler 2002	rat
HNF1 α	AKR1C4		ozeki 2002	ozeki 2002	human
HNF1 α	cldn2			sakaguchi 2002	human, mouse
HNF1 α	fgfr4		shah 2002	shah 2002	human
HNF1 α	lgf1			yang 2002	human/rat
HNF1 α	mif			yang 2002	human/rat
HNF1 α	Serpina1	Soutoglou 2002			human
HNF1 α	c1		zahedi 2002		human

Fig. 16

Name	RefSeq	Name	RefSeq	Name	RefSeq	Name	RefSeq	Name	RefSeq
A1BG	NM_130786	DHFR	NM_000791	GSS	NM_000178	ORC1L	NM_004153	UGT1A1	NM_000463
AASS	NM_005763	DKFZP434J037	NM_030952	H3FF	NM_003533	PABPC1	NM_002568	UGT2B11	NM_001073
ABCA8	NM_007168	DKFZP564O0523	NM_032120	H4FK	NM_003546	PCDHA12	NM_018903	UGT2B15	NM_001076
ABCB11	NM_003742	DKFZP586A0522	NM_014033	HABP2	NM_004132	PCK1	NM_002591	URKL1	NM_017859
ABCC2	NM_000392	DXF68S1E	NM_012080	HBP1	NM_012257	PHTF1	NM_006608	VCP	NM_007126
ABL2	NM_007314	E2F1	NM_005225	HCAP-G	NM_022346	PIK4CB	NM_002651	VTN	NM_000638
ACVR1	NM_001105	E2F1	NM_005225	HESX1	NM_003865	PLGL	NM_002665	WDR12	NM_018256
ADH1A	NM_000667	EIF4A1	NM_001416	HIVEP3	NM_024503	POLR2D	NM_004805	WDR5B	NM_019069
ADH1B	NM_000668	EIF4E	NM_001968	HMGCR	NM_000859	POLS	NM_006999		
AF038169	NM_013310	ELOVL1	NM_016031	HNF4a7	AF509467	PON1	NM_000446		
AGTR1	NM_000685	EPHA1	NM_005232	HNMT	NM_006895	PPFIA1	NM_003626		
AKR1C4	NM_001818	F11	NM_019559	HNRPR	NM_005826	PPP2R5A	NM_006243		
ALDH3A1	NM_000691	F9	NM_000133	HSD17B4	NM_000414	PRO1855	NM_018509		
ALDH5A1	NM_001080	FABP5	NM_001444	HSP105B	NM_006644	PSMA1	NM_002786		
AMBP	NM_001633	FACTP140	NM_007192	HSPA1B	NM_005346	PSMB1	NM_002793		
AMT	NM_000481	FADS3	NM_021727	HTR2B	NM_000867	PTPRR	NM_002849		
APCS	NM_001639	FLJ10209	NM_018026	IF	NM_000204	REA	NM_007273		
APOH	NM_000042	FLJ10407	NM_018087	INSM2	NM_032594	RING1	NM_002931		
ASPA	NM_000049	FLJ10415	NM_018089	IRF3	NM_001571	RNF20	NM_019592		
BCAR1	NM_014567	FLJ10578	NM_018144	IRF6	NM_006147	RPL35	NM_007209		
BCKDHA	NM_000709	FLJ10650	NM_018168	ITGAV	NM_002210	RPL37AP1	NG_000988		
BF	NM_001710	FLJ11029	NM_018304	ITIH1	NM_002215	RPLP1	NM_001003		
BM039	NM_018455	FLJ11105	NM_018323	JKK	NM_016281	RPS6KA5	NM_004755		
BNIP3L	NM_004331	FLJ11301	NM_018385	KIAA0806	NM_014813	RRP46	NM_020158		
BTN3A2	NM_007047	FLJ11726	NM_024971	KIAA0872	NM_014940	SART3	NM_014706		
C1S	NM_001734	FLJ11773	NM_021934	KIAA1056	NM_014894	SAS10	NM_020368		
C2	NM_000063	FLJ12552	NM_022832	KLF3	NM_016531	SCYB13	NM_006419		
C20orf188	NM_015638	FLJ12770	NM_032174	LIMK1	NM_016735	SEC10L1	NM_006544		
C8B	NM_000066	FLJ12910	NM_024573	LOC51060	NM_015913	SERPING1	NM_000062		
C8G	NM_000606	FLJ13798	NM_024773	LOC51074	NM_015957	SERPINI1	NM_005025		
CACNA1D	NM_000720	FLJ14153	NM_022736	LOC51287	NM_016565	SILV	NM_006928		
CASP2	NM_032982	FLJ20084	NM_017659	LOC51633	NM_016023	SLC1A3	NM_004172		
CCT8	NM_006585	FLJ20156	NM_017891	LOC51646	NM_016061	SLC25A13	NM_014251		
CDC25A	NM_001789	FLJ20422	NM_017814	LOC56906	NM_020147	SLC7A9	NM_014270		
CDC2L5	NM_003718	FLJ20627	NM_017909	LOC81558	NM_030802	SMARCC1	NM_003074		
CDK2	NM_001798	FLJ20671	NM_017924	LOH11CR2A	NM_014622	SMCY	NM_004653		
CDSN	NM_001264	FLJ20772	NM_017956	M17S2	NM_031858	SNRPD2	NM_004597		
CFL1	NM_005507	FLJ21934	NM_024743	MAP2K5	NM_002757	SNW1	NM_012245		
CH25H	NM_003956	FLJ21963	NM_024560	MGC10500	NM_031477	SNX3	NM_003795		
CLCN3	NM_001829	FLJ22169	NM_024085	MGC13053	NM_032710	SPG4	NM_014946		
CLDN2	NM_020384	FLJ22557	NM_024713	MGC16169	NM_033115	SPINK1	NM_003122		
CLLD8	NM_031915	FLJ23071	NM_025192	MGC16386	NM_080668	SPP2	NM_006944		
COL5A1	NM_000093	FLJ23263	NM_025115	MGC4189	NM_032308	SRF	NM_003131		
COL5A3	NM_015719	FLJ23375	NM_024956	MGST3	NM_004528	STMN2	NM_007029		
COPB2	NM_004766	FLJ23499	NM_022761	MN1	NM_002430	TAF2GL	NG_001012		
COPS7A	NM_016319	FLJ23598	NM_024783	NEK6	NM_014397	TAT	NM_000353		
CRADD	NM_003805	FXYP7	NM_022006	NFKBIA	NM_020529	TBX2	NM_005994		
CRI1	NM_014335	G6PC	NM_000151	NFKBIA	NM_020529	TCEB3	NM_003198		
CRP	NM_000567	GABPA	NM_002040	NFKBIA	NM_020529	TM4SF4	NM_004617		
CSN2	NM_001891	GAL3ST2	NM_033036	NOLC1	NM_004741	TMF1	NM_007114		
CYGB	NM_134268	GBF1	NM_004193	NR1I2	NM_022002	TMOD2	NM_014548		
CYP3A43	NM_022820	GJB1	NM_000166	NTF2	NM_005796	TNFRSF6	NM_000043		
CYP51	NM_000786	GRB2	NM_002086	OAT	NM_000274	TNFSF10	NM_003810		
D13S106E	NM_005800	GRO1	NM_001511	OAZ2	NM_002537	TOMM70A	NM_014820		
DDB2	NM_000107	GRO3	NM_002090	OGFR	NM_007346	TSG101	NM_006292		

Fig. 17

Name	RefSeq	Name	RefSeq	Name	RefSeq	Name	RefSeq
AASS	NM_005763	FLJ11271	NM_018373	JK	NM_016281	SEMA6A	NM_020796
ABCB8	NM_007188	FLJ11301	NM_018385	KIAA0660	NM_012297	SERPIN8	NM_002640
ACPP	NM_001099	FLJ11773	NM_021934	KIAA0712	NM_014715	SERPING1	NM_000062
ACVR1	NM_001105	FLJ12770	NM_032174	KIAA0806	NM_014813	SERPINI1	NM_005025
ADH1A	NM_000667	FLJ12910	NM_024573	KIAA0872	NM_014940	SH3BGR1	NM_003022
AF038169	NM_013310	FLJ13220	NM_021927	KIAA1056	NM_014894	SLC1A3	NM_004172
AF15Q14	NM_020380	FLJ13798	NM_024773	KRTAP1.1	NM_030967	SNRPD2	NM_004597
AGT	NM_000029	FLJ13955	NM_024759	LAMC2	NM_018891	SNW1	NM_012245
AMBP	NM_001633	FLJ14153	NM_022736	LBC	NM_006738	SPG4	NM_014946
AMT	NM_000481	FLJ14486	NM_032792	LOC51060	NM_015913	SPINK1	NM_003122
APCS	NM_001639	FLJ20084	NM_017659	LOC51287	NM_016565	TEGT	NM_003217
APOH	NM_000042	FLJ20156	NM_017691	LOC51633	NM_016023	TMF1	NM_007114
ARL1	NM_001177	FLJ20422	NM_017814	LOC56906	NM_020147	TNFRSF6	NM_000043
BBP	NM_032027	FLJ20627	NM_017909	LOC81558	NM_030802	TNFRSF6	NM_000043
BCKDHA	NM_000709	FLJ20643	NM_017916	LOH11CR2A	NM_014622	TNFRSF6	NM_000043
BF	NM_001710	FLJ20671	NM_017924	LUC7A	NM_016424	TNFRSF6	NM_000043
BTN3A2	NM_007047	FLJ20772	NM_017956	MDH1	NM_005917	TNFSF10	NM_003810
C1S	NM_001734	FLJ21272	NM_025032	MDS029	NM_018464	TOMM70A	NM_014820
C20orf188	NM_015638	FLJ21934	NM_024743	MEIS1	NM_002398	UGT2B15	NM_001076
C2F	NM_006331	FLJ21963	NM_024560	MGC13040	NM_032930	UGT2B17	NM_001077
C8orf4	NM_020130	FLJ22169	NM_024085	MGC13053	NM_032710	VCP	NM_007126
CCT8	NM_006585	FLJ23263	NM_025115	MGC19595	NM_033415	VTN	NM_000638
CCD2L5	NM_003718	FLJ23375	NM_024956	MGC3020	NM_024048	WDR12	NM_018256
CH25H	NM_003956	GABARAPL1	NM_031412	MGC3413	NM_032678	ZNF317	NM_020933
CIR	NM_004882	GABPA	NM_002040	MGC4189	NM_032308		
CLCN4	NM_001830	GCP3	NM_006322	MGST3	NM_004528		
CLDN2	NM_020384	GJB1	NM_000166	MTERF	NM_006980		
CLLD8	NM_031915	GLA	NM_000169	NET-6	NM_014399		
CLNS1A	NM_001293	GRB2	NM_002086	NOLC1	NM_004741		
CLONE24922	NM_015679	GRO1	NM_001511	NOVA1	NM_006489		
CMG1	NM_025103	GRO3	NM_002090	NR0B2	NM_021969		
COPB2	NM_004766	GSS	NM_000178	NUDT2	NM_001161		
COPS7A	NM_016319	GSTA4	NM_001512	OGFR	NM_007346		
COX4I1	NM_001861	GTF2E1	NM_005513	ORC1L	NM_004153		
COX7A2L	NM_004718	H4FA	NM_003538	PAPA-1	NM_031288		
CRI1	NM_014335	H4FH	NM_003543	PEX6	NM_000287		
GSN2	NM_001891	HABP2	NM_004132	PMAIP1	NM_021127		
CYP3A43	NM_022820	HASJ4442	NM_017528	PPFIA1	NM_003626		
DKFZp761D221	NM_032291	HBOA	NM_007067	PPFIBP1	NM_003622		
DKFZp761J139	NM_032280	HBP1	NM_012257	PPP1R3D	NM_006242		
EED	NM_003797	HLA-G	NM_002127	PSMA1	NM_002786		
EGR2	NM_000399	HMG2	NM_002129	PSMB1	NM_002793		
EHD4	NM_014599	HNF4a7	AF509467	PTPRN2	NM_002847		
EHF	NM_012153	HNRPA2B1	NM_031243	REA	NM_007273		
EIF4E	NM_001968	HNRPR	NM_005826	RECK	NM_021111		
F11	NM_019559	HSD17B4	NM_000414	RIG-I	NM_014314		
F2RL2	NM_004101	HSN44A4A	NM_015372	RPC32	NM_006467		
FABP5	NM_001444	HSP105B	NM_006644	RPL36P1	NG_000983		
FER1L3	NM_133337	HSPA1B	NM_005346	RPS6KA5	NM_004755		
FLJ10342	NM_018064	HSPC125	NM_014165	RRP46	NM_020158		
FLJ10407	NM_018087	HT007	NM_018480	SAMHD1	NM_015474		
FLJ10415	NM_018089	HTR2B	NM_000867	SART3	NM_014706		
FLJ10482	NM_018107	humNRDR	NM_021004	SAS10	NM_020368		
FLJ10650	NM_018168	IGSF3	NM_001542	SCYA28	NM_019846		
FLJ11029	NM_018304	IRF3	NM_001571	SEC10L1	NM_006544		

[illegible]

Gene Name	RefSeq	Gene Name	RefSeq	Gene Name	RefSeq	Gene Name	RefSeq	Gene Name	RefSeq	Gene Name	RefSeq	Gene Name	RefSeq	Gene Name	RefSeq
GTF2P1	NM_005316	IGF1	NM_000518	LDC51015	NM_016048	IGTAP2	NM_056013	MPRL49	NM_004927	OSGEP	NM_017807	PPPIR12B	NM_032106	PPPIR12B	NM_032106
GTF3P3	NM_032620	IGFBP1	NM_000396	LDC51016	NM_016049	METAB48	NM_054038	MPRL51	NM_016497	OSMR	NM_003999	PPPIR15B	NM_032833	PPPIR15B	NM_032833
GTF3P3	NM_032620	IGFBP1	NM_000396	LDC51017	NM_016072	MGC13379	NM_016499	MRSP11	NM_022839	OTC	NM_000531	PPPIR3C	NM_024607	PPPIR3C	NM_024607
H2AFS	NM_005896	IL1R1	NM_005653	LDC51022	NM_016072	MGC10433	NM_024321	MRSP14	NM_022100	OTC	NM_000531	PPPIR3C	NM_024607	PPPIR3C	NM_024607
H2AFG	NM_021065	IL1RA	NM_002182	LDC51054	NM_015989	MGC10702	NM_023653	MRSP18	NM_016065	P115	NM_037115	PPPIR3D	NM_026422	PPPIR3D	NM_026422
H2AFO	NM_003516	IL22R	NM_021258	LDC51060	NM_015913	MGC10823	NM_021437	MRSP18B	NM_016497	P21UA58	NM_003389	PPPIR3E	NM_027151	PPPIR3E	NM_027151
H2BFA	NM_003518	IL2RB	NM_000878	LDC51064	NM_015917	MGC10624	NM_030571	MRSP18C	NM_016067	P21UA58	NM_003389	PPPIR3E	NM_027151	PPPIR3E	NM_027151
H2BFB	NM_021063	IL6ST	NM_002184	LDC51074	NM_015957	MGC10904	NM_032303	MRSP21	NM_018997	P21UA59	NM_003389	PPPIR41	NM_031434	PPPIR41	NM_031434
H2BFF	NM_021062	IL6MT	NM_006839	LDC51091	NM_015955	MGC10960	NM_032353	MRSP30	NM_016540	P21UA59	NM_003389	PPPIR41	NM_031434	PPPIR41	NM_031434
H2BFG	NM_003522	IL6NL	NM_005799	LDC51094	NM_016001	MGC10974	NM_032308	MRSP35	NM_018211	P29	NM_015484	PPPIR41	NM_031434	PPPIR41	NM_031434
H2BEG	NM_015722	INHBC	NM_005358	LDC51106	NM_016014	MGC10999	NM_032307	MRSP36	NM_033281	P29Y2	NM_002654	PPPIR41	NM_031434	PPPIR41	NM_031434
H2BEG	NM_015722	INHBC	NM_005358	LDC51107	NM_016022	MGC10314	NM_031453	MRSP57	NM_015971	PABPC1	NM_002658	PPPIR41	NM_031434	PPPIR41	NM_031434
H2BEG	NM_015722	INHBC	NM_005358	LDC51108	NM_016022	MGC11286	NM_032322	MRSL2	NM_020862	PABPC1	NM_002658	PPPIR41	NM_031434	PPPIR41	NM_031434
H2BEG	NM_015722	INHBC	NM_005358	LDC51109	NM_016022	MGC11271	NM_032322	MRSL2	NM_020862	PABPC1	NM_002658	PPPIR41	NM_031434	PPPIR41	NM_031434
H2BEG	NM_015722	INHBC	NM_005358	LDC51110	NM_016022	MGC11271	NM_032322	MRSL2	NM_020862	PABPC1	NM_002658	PPPIR41	NM_031434	PPPIR41	NM_031434
H2BEG	NM_015722	INHBC	NM_005358	LDC51111	NM_016022	MGC11271	NM_032322	MRSL2	NM_020862	PABPC1	NM_002658	PPPIR41	NM_031434	PPPIR41	NM_031434
H2BEG	NM_015722	INHBC	NM_005358	LDC51112	NM_016022	MGC11271	NM_032322	MRSL2	NM_020862	PABPC1	NM_002658	PPPIR41	NM_031434	PPPIR41	NM_031434
H2BEG	NM_015722	INHBC	NM_005358	LDC51113	NM_016022	MGC11271	NM_032322	MRSL2	NM_020862	PABPC1	NM_002658	PPPIR41	NM_031434	PPPIR41	NM_031434
H2BEG	NM_015722	INHBC	NM_005358	LDC51114	NM_016022	MGC11271	NM_032322	MRSL2	NM_020862	PABPC1	NM_002658	PPPIR41	NM_031434	PPPIR41	NM_031434
H2BEG	NM_015722	INHBC	NM_005358	LDC51115	NM_016022	MGC11271	NM_032322	MRSL2	NM_020862	PABPC1	NM_002658	PPPIR41	NM_031434	PPPIR41	NM_031434
H2BEG	NM_015722	INHBC	NM_005358	LDC51116	NM_016022	MGC11271	NM_032322	MRSL2	NM_020862	PABPC1	NM_002658	PPPIR41	NM_031434	PPPIR41	NM_031434
H2BEG	NM_015722	INHBC	NM_005358	LDC51117	NM_016022	MGC11271	NM_032322	MRSL2	NM						

Fig. 18C

Gene Name	RefSeq	Gene Name	RefSeq	Gene Name	RefSeq	Gene Name	RefSeq
RNPC2	NM_004902	SLC25A13	NM_014251	TDRKH	NM_006882	VPS45A	NM_007259
RNPEPL1	NM_018226	SLC25A18	NM_031481	TEAD3	NM_003214	VTN	NM_000638
ROCK1	NM_005406	SLC25A5	NM_001152	TED	NM_015686	WASF3	NM_006646
RORC	NM_005060	SLC26A1	NM_022042	TEF	NM_003216	WASL	NM_003941
RPC32	NM_005467	SLC2A8	NM_014580	TEGT	NM_003217	WBP4	NM_007187
RPL18	NM_000979	SLC31A1	NM_001859	TESK2	NM_007170	WDF2	NM_052950
RPL31	NM_000993	SLC35A2	NM_005660	TF	NM_001063	WDR10	NM_052985
RPL37AP1	NG_000988	SLC35A3	NM_012243	THPO	NM_000460	WDR12	NM_018256
RPL5	NM_000969	SLC38A1	NM_030574	THTP	NM_024328	WDR13	NM_017883
RPL7	NM_000971	SLC38A4	NM_018018	TIA1	NM_022037	XDH	NM_000379
RPLP1	NM_001003	SLC39A1	NM_014437	TIMM17A	NM_006335	XPA	NM_000380
RPS16	NM_001020	SLC5A3	NM_006933	TIMM17B	NM_005834	XPC	NM_004628
RPS19	NM_001022	SLC7A2	NM_003046	TIMM23	NM_006327	XPR1	NM_004736
RPS27A	NM_002954	SLC7A9	NM_014270	TIMM9	NM_012460	XRCC5	NM_021141
RPS3A	NM_001008	SLP1	NM_003064	TLH29	NM_032036	YKT6	NM_006555
RPS6KA5	NM_004755	SMAC	NM_019887	TLN1	NM_006289	YWHAB	NM_003404
RPS6KB1	NM_003161	SMAP	NM_006896	TM4SF4	NM_004617	ZAN	NM_003386
RQCD1	NM_005444	SMARCA5	NM_003601	TM9SF2	NM_004800	ZBRK1	NM_021632
RSHL1	NM_030785	SMARCE1	NM_003079	TMEM7	NM_031440	ZF5128	NM_014347
RSP3	NM_031924	SMC2L1	NM_006444	TMF1	NM_007114	ZFP95	NM_014569
RSU1	NM_012425	SMPD1	NM_000543	TMOD2	NM_014548	ZK1	NM_005815
RTC01	NM_003729	SNAI2	NM_003068	TMP21	NM_006827	ZNF133	NM_003434
RTP801	NM_019058	SNAP23	NM_003825	TNFAIP1	NM_021137	ZNF144	NM_007144
RUVBL2	NM_006666	SNAPC1	NM_003082	TNFRSF1B	NM_002546	ZNF146	NM_007145
S100A9	NM_002955	SNK	NM_006622	TNFRSF6	NM_000043	ZNF147	NM_005082
SA1	NM_000331	SNRPA	NM_004596	TNFRSF6	NM_000043	ZNF155	NM_003445
SA1	NM_000331	SNRPD3	NM_004175	TNFRSF6	NM_000043	ZNF183	NM_006978
SA1	NM_000331	SNRPF	NM_003095	TNFRSF6	NM_000043	ZNF192	NM_006298
SA1	NM_000331	SNW1	NM_012245	TNFSF13	NM_003808	ZNF207	NM_003457
SA2	NM_030754	SNX1	NM_003099	TNS	NM_022648	ZNF214	NM_013249
SAC	NM_018417	SNX17	NM_014748	TOM1	NM_005488	ZNF22	NM_006963
SAD1	NM_006590	SNX5	NM_003795	TOMM70A	NM_014820	ZNF221	NM_013359
SAMHD1	NM_015474	SOD1	NM_000454	TP53TG1	NM_007233	ZNF222	NM_013360
SAP18	NM_005870	SORCS3	NM_014978	TPP2	NM_003291	ZNF224	NM_013398
SAS10	NM_020358	SOX10	NM_006941	TPT	NM_014317	ZNF225	NM_013362
SC4MOL	NM_006745	SP2	NM_138406	TRA1	NM_003299	ZNF226	NM_016444
SCA2	NM_002973	SPATA2	NM_006038	TRAF6	NM_004620	ZNF230	NM_006300
SCAND1	NM_033630	SPATA6	NM_019073	TRAP150	NM_005119	ZNF237	NM_014242
SCD	NM_005063	SPC18	NM_014300	TRIM15	NM_033229	ZNF281	NM_012482
SCYA14	NM_032962	SPOCK	NM_004598	TRIM26	NM_003449	ZNF302	NM_018443
SCYA15	NM_032964	SPP2	NM_006944	TRIM31	NM_052816	ZNF361	NM_018555
SCYA16	NM_004590	SQRDL	NM_021199	TRIM34	NM_130389	ZNF9	NM_003418
SCYE1	NM_004757	SREBF2	NM_004599	TRIM4	NM_033017	ZNF-U69274	NM_014415
SDC1	NM_002997	SRP54	NM_003136	TRIP11	NM_004239	ZNRD1	NM_014596
SDCCAG10	NM_005869	SRP68	NM_014230	TRN-SR	NM_012470	ZnTL2	NM_133496
SDCCAG28	NM_006645	SRPR	NM_003139	TRPC5	NM_012471		
SDFR1	NM_012428	SSA2	NM_004600	TRPS1	NM_014112		
SEC10L1	NM_006544	SSAT2	NM_133491	TSG101	NM_006292		
SEC23A	NM_006364	SSSCA1	NM_006396	TSLRP	NM_012472		
SEC24D	NM_014822	SSTR1	NM_001049	TTY14	NM_031932		
SEC8B18	NM_006808	STAF42	NM_053053	TUBB5	NM_006087		
SEL1L	NM_005085	STAF65(gamma)	NM_014860	TUFT1	NM_020127		
SEMA3C	NM_006379	STAM	NM_003473	TXNIP	NM_006472		
SEMA6C	NM_030913	STAM2	NM_005843	TXNL	NM_004786		
SEMA7A	NM_003612	STARD7	NM_020151	TXNRD1	NM_003330		
SENP1	NM_014554	STAT1	NM_007315	TYMS	NM_001071		
SEPX1	NM_016332	STAU2	NM_014393	U2AF1	NM_006758		
SERPINA1	NM_000295	STCH	NM_006948	U3-55K	NM_004704		
SERPINA10	NM_016186	STIM1	NM_003156	US-116KD	NM_004247		
SERPINA5	NM_000624	STK19	NM_004197	UBE2B	NM_003337		
SERPINA6	NM_001756	STK2	NM_003157	UBE2D3	NM_003340		
SERPINC1	NM_000488	STOML1	NM_004809	UBE2M	NM_003969		
SERPIND1	NM_000185	STRAIT11499	NM_021242	UBP1	NM_014517		
SERPINE1	NM_000602	STX18	NM_016930	UBQLN1	NM_053067		
SERPING1	NM_000062	SUCLA2	NM_003850	UBQLN2	NM_013444		
SERPINI1	NM_005025	SUCLG1	NM_003849	UCH37	NM_015984		
SES2	NM_031459	SUDD	NM_003831	UCHL3	NM_006002		
SF3A3	NM_006802	SULT1A1	NM_001055	UGDH	NM_003359		
SF3B2	NM_006842	SULT2A1	NM_003167	UGT2B11	NM_001073		
SFRS11	NM_004768	SUOX	NM_000456	UGT2B15	NM_001076		
SFRS5	NM_006925	SUPT3H	NM_003599	UGTREL1	NM_005827		
SFRS8	NM_004592	SUPT5H	NM_003169	UGTREL7	NM_015139		
SGK	NM_005627	SUPV3L1	NM_003171	ULBP3	NM_024518		
SGK2	NM_016276	SYN3	NM_133632	UPB1	NM_016327		
SGT1	NM_006704	SYTL4	NM_080737	UQCRC2	NM_003366		
SH2D3C	NM_005489	SZF1	NM_016089	URKL1	NM_017859		
SH3BGR12	NM_031469	TADA3L	NM_133480	UROD	NM_000374		
SILV	NM_006928	TAF2GL	NG_001012	UROS	NM_000375		
SIX2	NM_016932	TAGLN2	NM_003564	USP1	NM_003368		
SKB1	NM_005109	TARS	NM_003191	USP15	NM_006313		
SKD1	NM_004869	TAT	NM_000353	USP2	NM_004205		
SKRP1	NM_080876	TCF1	NM_000545	UST	NM_004182		
SLC10A1	NM_003049	TCF12	NM_003205	VAMP1	NM_014231		
SLC17A2	NM_005835	TCF19	NM_007109	VAMP5	NM_006634		
SLC17A5	NM_012434	TCF21	NM_003206	VDAC1	NM_003374		
SLC19A3	NM_025243	TCF7L2	NM_030756	VDAC2	NM_003375		
SLC22A1LS	NM_007105	TCIRG1	NM_005019	VEGFC	NM_005429		
SLC22A2A	NM_021977	TCOP1	NM_000356	VEZATIN	NM_017599		
SLC22A7	NM_006672	TCP1	NM_030752	VMP1	NM_030938		
				VPS29	NM_016226		

[illegible]

Fig. 19B

Gene Name	RefSeq	Gene Name	RefSeq	Gene Name	RefSeq	Gene Name	RefSeq	Gene Name	RefSeq	Gene Name	RefSeq	Gene Name	RefSeq	Gene Name	RefSeq
IGSF8	NM_052868	LOC51075	NM_015959	MFAP1	NM_005926	MRPL33	NM_004891	NTE	NM_006702	PPP1R15B	NM_032833	RPL7	NM_000971		
IMAGE145052	NM_014287	LOC51076	NM_015960	MGC10433	NM_024321	MRPL43	NM_032112	NTF2	NM_005796	PPP2R5B	NM_006244	RPLP0L	NM_016183		
IMAGE345520	NM_024005	LOC51077	NM_015962	MGC10500	NM_031477	MRPL44	NM_022915	NUCB1	NM_008184	PPP6C	NM_002721	Rpl01-2	NM_019014		
IMBT	NM_008939	LOC51094	NM_015999	MGC10702	NM_032663	MRPL46	NM_022163	NUDT2	NM_001161	PRCC	NM_005973	RP514	NM_005617		
IMP13	NM_014652	LOC51095	NM_016001	MGC10924	NM_033571	MRPL48	NM_016055	NUDT5	NM_014142	PRDM5	NM_016659	RP518	NM_001020		
INCEP	NM_020238	LOC51104	NM_016014	MGC10974	NM_033306	MRPL51	NM_016497	NUDT6	NM_007083	PRDX5	NM_012094	RP519	NM_022551		
ING3	NM_019071	LOC51107	NM_016022	MGC10999	NM_032307	MRPL53	NM_053050	NUPI07	NM_024001	PRKAB1	NM_005253	RP519	NM_001022		
ING4	NM_016162	LOC51117	NM_016035	MGC11102	NM_032325	MRPS11	NM_022839	NUP54	NM_017426	PRKCA8P	NM_012407	RP520	NM_001023		
INVS	NM_014425	LOC51118	NM_016037	MGC11115	NM_032310	MRPS12	NM_021107	NUP62	NM_012346	PRKCE	NM_005400	RP521	NM_001024		
IRS4	NM_003604	LOC51142	NM_016139	MGC11266	NM_024322	MRPS14	NM_022100	NVL	NM_025533	PROZ389	NM_025230	RP525	NM_001028		
ITGA6	NM_000210	LOC51174	NM_016261	MGC11127	NM_033549	MRPS15	NM_031280	NYD-SP11	NM_031951	PRP18	NM_003675	RP527A	NM_002954		
ITGA9	NM_002207	LOC51189	NM_016304	MGC11279	NM_024326	MRPS16	NM_016065	NY-REN-41	NM_006054	PRPF31	NM_015629	RP528	NM_001031		
ITGB3BP	NM_014288	LOC51202	NM_016355	MGC11296	NM_023352	MRPS18B	NM_014046	OBTP	NM_013397	PRRG2	NM_000991	RP53	NM_001005		
ITM1	NM_002219	LOC51204	NM_016360	MGC11352	NM_033927	MRPS18C	NM_016057	OGFR	NM_007346	PRSS25	NM_013247	RP53A	NM_001008		
JM4	NM_007213	LOC51205	NM_016361	MGC12943	NM_033317	MRPS21	NM_018997	OPA1	NM_015580	PSCD2	NM_004225	RP55	NM_001009		
JTB	NM_006594	LOC51231	NM_016440	MGC12981	NM_032357	MRPS23	NM_016070	OPA3	NM_025136	PSMA1	NM_002786	RP56	NM_001010		
KARS	NM_005548	LOC51246	NM_016479	MGC13102	NM_032323	MRPS27	NM_015094	ORC1L	NM_004153	PSMA2	NM_002787	RP56K45	NM_004755		
KBRAS1	NM_020345	LOC51287	NM_016565	MGC13114	NM_032366	MRPS28	NM_014018	ORC3L	NM_012381	PSMA3	NM_002788	RP56K61	NM_003151		
KCNQ5	NM_019842	LOC51290	NM_016570	MGC13138	NM_033410	MRPS30	NM_016640	OSBP	NM_002556	PSMA5	NM_002790	RP56K61	NM_003151		
KEO4	NM_006459	LOC51292	NM_016576	MGC13159	NM_032027	MRPS35	NM_021821	OSBP11	NM_022776	PSMB1	NM_002793	RP56K61	NM_003151		
KIAA0028	NM_015340	LOC51300	NM_016589	MGC1346	NM_032758	MRPS7	NM_015971	OSCAR	NM_013077	PSMB5	NM_002797	RRP4	NM_014285		
KIAA0057	NM_012268	LOC51326	NM_016632	MGC14126	NM_032859	MSMB	NM_002443	OSGEP	NM_017807	PSMB7	NM_002759	RRP46	NM_020158		
KIAA0092	NM_014679	LOC51329	NM_016638	MGC14151	NM_032356	MSPT28	NM_031554	P125	NM_007190	PSMCA	NM_005930	RSU1	NM_012425		
KIAA0102	NM_014752	LOC51566	NM_015921	MGC14208	NM_032901	MTERF	NM_006990	P16-2	NM_018689	PSMD1	NM_002801	RXRB	NM_021976		
KIAA0105	NM_004906	LOC51604	NM_015937	MGC14241	NM_032907	MTR1	NM_005955	P29	NM_015484	PSMD10	NM_002814	SACM2L	NM_002553		
KIAA0154	NM_014739	LOC51605	NM_015939	MGC14595	NM_032334	MTHFR1	NM_005956	P5326	NM_031450	PSMD4	NM_002810	SAD1	NM_005950		
KIAA0196	NM_014846	LOC51626	NM_016008	MGC14697	NM_032747	MTR4	NM_004687	PACE	NM_002569	PSMD7	NM_002811	SAP18	NM_005870		
KIAA0255	NM_014742	LOC51631	NM_016019	MGC14788	NM_006050	MTR	NM_000254	PACE4	NM_002570	PSMD8	NM_002812	SART3	NM_014708		
KIAA0258	NM_014785	LOC51633	NM_016023	MGC14836	NM_033412	MTRF1	NM_004294	PAFAH2	NM_000437	PSME3	NM_005789	SAS10	NM_020368		
KIAA0274	NM_014845	LOC51644	NM_016057	MGC15677	NM_032878	mTRF	NM_016020	PAIRBP1	NM_015640	PTD009	NM_016146	SBP2	NM_024077		
KIAA0317	NM_014821	LOC51651	NM_016077	MGC15169	NM_033115	MUT	NM_000255	PANX2	NM_025839	PTD012	NM_014039	SCDFG-B	NM_025208		
KIAA0372	NM_014639	LOC51657	NM_016086	MGC16386	NM_006068	MUTYH	NM_012222	PAPA-1	NM_031298	PTD013	NM_015552	SCML1	NM_006748		
KIAA0391	NM_014672	LOC51691	NM_016200	MGC16733	NM_033547	MXI1	NM_005952	PARV9	NM_013327	PTD015	NM_014016	SCY1	NM_004757		
KIAA0416	NM_015564	LOC54516	NM_019041	MGC17347	NM_033833	MYCBP	NM_012333	PAWR	NM_002583	PTK7	NM_002821	SDCCAG10	NM_005889		
KIAA0419	NM_014711	LOC54543	NM_019059	MGC19595	NM_033415	MYL6	NM_079424	PAX1	NM_005192	PTPH13	NM_006254	SDCCAG28	NM_006455		
KIAA0426	NM_014724	LOC55815	NM_018430	MGC20486	NM_052844	NAG	NM_015909	pcnp	NM_020357	PWP1	NM_007082	SDF2	NM_006923		
KIAA0433	NM_015216	LOC55954	NM_019103	MGC2404	NM_032360	NAGK	NM_017567	PCQAP	NM_015889	R3HDM	NM_015381	SDFR1	NM_012428		
KIAA0438	NM_014819	LOC55851	NM_020154	MGC2408	NM_032331	NAKAP95	NM_014371	PCYT1A	NM_005017	RA410	NM_016106	SDHC	NM_003001		
KIAA0537	NM_014840	LOC55902	NM_020143	MGC24447	NM_032878	NAPA	NM_003827	PDCD10	NM_007217	RAB11A	NM_004663	SEC10L1	NM_006544		
KIAA0547	NM_014783	LOC55993	NM_020243	MGC2474	NM_032671	NBP	NM_023533	PDE4DIP	NM_014844	RAB18	NM_021252	SEC22L1	NM_004832		
KIAA0570	NM_014877	LOC57019	NM_020313	MGC2477	NM_024059	NBR2	NM_005821	PDE5D	NM_005951	RAB19	NM_020561	SEC3	NM_014651		
KIAA0582	NM_014852	LOC57107	NM_020381	MGC2488	NM_024039	NCBP1	NM_002486	PDE9A	NM_020806	RAB2	NM_002865	SEC8B	NM_005818		
KIAA0710	NM_014871	LOC57109	NM_020385	MGC2508	NM_024327	NCBP2	NM_007362	PEAS	NM_057161	RAB30	NM_014488	SEC8	NM_021607		
KIAA0756	NM_014805	LOC57147	NM_020423	MGC2560	NM_031452	NCOA4	NM_005437	PEF	NM_012392	RAB6KIFL	NM_005733	SEDL	NM_015890		
KIAA0765	NM_025010	LOC58929	NM_022098	MGC2650	NM_024108	NDUFA1	NM_004541	PEMT	NM_007169	RAB7	NM_004637	SEL1L	NM_005065		
KIAA0806	NM_014813	LOC81034	NM_030780	MGC2855	NM_024339	NDUFA3	NM_004542	PET112L	NM_004564	RABAC1	NM_006423	SEN1P	NM_014554		
KIAA0872	NM_014940	LOC81558	NM_030802	MGC2747	NM_024104	NDUFA4	NM_002489	PEX11B	NM_003848	Rab4R	NM_017987	SERPINA4	NM_006215		
KIAA0907	NM_014949	LOC89553	NM_038343	MGC2840	NM_024078	NDUFA5	NM_005000	PEX12	NM_000286	RAD51	NM_013487	SERPINB3	NM_006919		
KIAA0950	NM_014929	LOC90346	NM_038351	MGC3121	NM_024031	NDUFA6	NM_002490	PEX13	NM_002616	RAGA	NM_006570	SERPINEB	NM_002640		
KIAA0971	NM_014929	LOC90878	NM_038361	MGC3123	NM_024107	NDUFA7	NM_005001	PEX16	NM_002617	RA2	NM_021785	SERPIN1	NM_006925		
KIAA1012	NM_014939	LOC90701	NM_033280	MGC3133	NM_031287	NDUFB3	NM_002491	PEX3	NM_000287	RAMP	NM_016448	SESS	NM_014549		
KIAA1017	NM_007216	LOC90799	NM_038363	MGC3180	NM_024041	NDUFB8	NM_002492	PFDN5	NM_002624	RANBP8	NM_006390	SETDB1	NM_012432		
KIAA1041	NM_014947	LOC92105	NM_038381	MGC3222	NM_024334	NDUFS1	NM_005005	PHACS	NM_032592	RANGAP1	NM_002883	SF3A3	NM_006802		
KIAA1068	NM_015332	LRP5	NM_020335	MGC3248	NM_032466	NDUFS3	NM_004551	PHB	NM_002634	RAP1	NM_007189	SF3B1	NM_012433		
KIAA1100	NM_014901	LRN1	NM_002319	MGC4054	NM_024341	NDUFS4	NM_002495	PHKB	NM_000293	RARG-1	NM_016167	SF3B2	NM_006842		
KIAA1108	NM_024820	LSA3	NM_014463	MGC4093	NM_030578	NDUFV1	NM_007103	PIGN	NM_012327	RASSF1	NM_007182	SF3B4	NM_005850		
KIAA1775	NM_033100	LSM4	NM_012321	MGC4161	NM_024303	NEDD8	NM_006156	PICP1	NM_022121	REAK	NM_021813	SFRS1	NM_006924		
KIF3B	NM_004799	LSM5	NM_012322	MGC4189	NM_032308	NEK7	NM_033494	PICP2	NM_022121	REBP4	NM_005610	SFRS11	NM_004768		
KIF9	NM_022342	LTA4H	NM_000895	MGC4251	NM_032376	NFATC2	NM_012340	PICP3	NM_022121	RBL1	NM_002895	SFRS2	NM_003016		
KLRF1	NM_016523	LYPLA2	NM_007260	MGC4308	NM_032359	NFE2L3	NM_003204	PIC3C3	NM_002647	RBL2	NM_005611	SFRS5	NM_006925		
KNSL7	NM_020242	LZTF1L	NM_020347	MGC4608	NM_024516	NFE2L3	NM_004289	PINK1	NM_032409	REM15	NM_022768	SFRS8	NM_004592		
KPTN	NM_007059	LZTR1	NM_006767	MGC4767	NM_032314	NFKB1B	NM_002503	PIP5K1A	NM_003657	RBM6	NM_006777	SGCE	NM_003910		
KRT10	NM_000421	M17S2	NM_031858	MGC4771	NM_032668	NFKB2	NM_002503	PIST	NM_020399	RBM7	NM_016090	SGT1	NM_006704		
LAPTM4A	NM_014713	M5A	NM_019852	MGC5302	NM_024089	NFKB3	NM_002503	PL6	NM_007024	RDBP	NM_002904	SH3BGR2L	NM_031469		
LCHT	NM_019015	M9	NM_013234	MGC5347	NM_024083	NFKB4	NM_002503	PLA2G2D	NM_012400	RDH5	NM_002995	SHH	NM_000193		
LCP	NM_014315	M9B	NM_007358	MGC5378	NM_032632	NFKB1L	NM_005007	PLA2G4B	NM_005090	REA	NM_007273	SIP	NM_014412		
LD1	NM_003853	MAGOH	NM_002070	MGC5469	NM_032356	NEFY	NM_006206	PLAA	NM_004253	RECCO	NM_002907	SIRT2	NM_003593		
LEPR	NM_002303	MAP2K5	NM_002757	MGC5509	NM_024093	NIMP	NM_032730	PLDN	NM_013388	RECOL5	NM_004259	SIRT3	NM_006109		
LGSM	NM_005606	MAP3K11	NM_002419	MGC5521	NM_024081	NKTR	NM_005385	PHE-1	NM_016147	REG1A	NM_002909	SKD1	NM_004889		
LHX6	NM_014368	MAP3K3	NM_002401	MGC5984	NM_033418	NLN	NM_020726	PMS2	NM_000535	REG1B	NM_006507	SKD3	NM_003913		
LIM	NM_006457	MAP3K7	NM_003188	MGC9740	NM_006056	NMA	NM_012342	PMS2L2	NM_005394	RENT1	NM_002911	SKI	NM_003036		
LMAS1	NM_004987	MAPK7	NM_002749	MGST3	NM_004528	NME1	NM_000269	PNAS-131	NM_031448	RFC3	NM_002915	SKP2	NM_032637		
LN47-C	NM_018352	MAPK8IP2	NM_012324	MID1	NM_000381	NME7	NM_013330	PNKP	NM_007254	RFL2	NM_005605	SLC1A5A	NM_004694		
LUC7	NM_015025	MAPK8IP3	NM_015133	MKRN1	NM_013446	NHOG1	NM_019082	PNMA1	NM_006029	RNF40	NM_014771	SLC25A19	NM_021734		
LVI1	NM_012319	NAT2A	NM_005911	MLH1	NM_000249	NOLA1	NM_019893	PODXL	NM_005397	RNF5	NM_005913	SLC22A8	NM_014580		
LOC113251	NM_052879	NBD4	NM_003925	NLN	NM_002418	NOL3	NM_004741	POL3	NM_017443	RNGT1	NM_003800	SLC31A1	NM_007859		
LOC113444	NM_052879	NBD4	NM_003925	NLN	NM										

Fig. 19C

Gene Name	RefSeq	Gene Name	RefSeq
SNRPD3	NM_004175	TXNL	NM_0041
SNRPF	NM_003095	U2AF1	NM_0067
SNW1	NM_012245	U5-100K	NM_0046
SNX1	NM_003099	U5-116KD	NM_0042
SNX11	NM_013323	UBE2M	NM_0035
SNX17	NM_014748	UBE2N	NM_0033
SNX5	NM_014426	UBE2V1	NM_0224
SON	NM_003103	UBQLN1	NM_0530
SOX17	NM_022454	UCH37	NM_0155
SOX9	NM_006346	UGTREL1	NM_0058
SP2	NM_138406	UMPS	NM_0003
SPATA2	NM_006038	UNRIP	NM_0071
SPC18	NM_014300	UPF3B	NM_0806
SPG4	NM_014946	UQCRC2	NM_0033
SPK	NM_004819	UQCRLH	NM_0060
SORDL	NM_021199	URKL1	NM_0178
SRP19	NM_003135	UROD	NM_0003
SRP54	NM_003136	UROS	NM_0003
SRP68	NM_014230	USF1	NM_0071
SSA2	NM_004600	USP5	NM_0032
SSBP1	NM_003143	UXT	NM_0041
SSFA2	NM_006751	V1RL1	NM_0206
SSR2	NM_003145	VEGFC	NM_0054
SSR3	NM_007107	VMP1	NM_0305
SSSCA1	NM_006396	VPS33A	NM_0228
SSTK	NM_032037	WARS2	NM_0155
SSTR4	NM_001052	WBP4	NM_0071
ST13	NM_003932	WDF2	NM_0525
STAF42	NM_053053	WDR12	NM_0162
STAF65 (oam)	NM_014860	WDR13	NM_0178
STAM	NM_003473	WHIP	NM_0201
STAM2	NM_005843	XPC	NM_0046
STCH	NM_006948	XPO1	NM_0034
STK19	NM_004197	XRCC4	NM_0225
STK24	NM_003576	XRCC5	NM_0211
STOML1	NM_004809	XRN2	NM_0122
STOML2	NM_013442	YR-29	NM_0146
STX18	NM_016930	YWHAB	NM_0034
SUCLG1	NM_003849	ZBRK1	NM_0216
SULT1A3	NM_003166	ZF5128	NM_0143
SULT1C1	NM_001056	ZFP37	NM_0034
SUPT5H	NM_003169	ZFP93	NM_0042
SUPV3L1	NM_003171	ZFP95	NM_0146
T54	NM_015898	ZNF133	NM_0034
TADA3L	NM_133480	ZNF134	NM_0034
TAF11	NM_005643	ZNF142	NM_0050
TAF6	NM_005641	ZNF146	NM_0071
TARBP2	NM_004178	ZNF155	NM_0034
TAX1BP1	NM_006024	ZNF175	NM_0071
TCERG1	NM_006708	ZNF183	NM_0065
TCF1	NM_000645	ZNF189	NM_0034
TCF2	NM_000458	ZNF192	NM_0062
TCF2	NM_000458	ZNF193	NM_0062
TCF2	NM_000458	ZNF207	NM_0034
TCOF1	NM_000356	ZNF214	NM_0132
TCP1	NM_030752	ZNF221	NM_0133
TDRKH	NM_006862	ZNF222	NM_0133
TEGT	NM_003217	ZNF224	NM_0133
TESK2	NM_007170	ZNF225	NM_0133
TFAP4	NM_003223	ZNF226	NM_0164
TFPT	NM_013342	ZNF230	NM_0065
TG737	NM_006531	ZNF264	NM_0034
TIMM23	NM_006327	ZNF265	NM_0054
TIMM9	NM_012460	ZNF277	NM_0216
TIP39	NM_012143	ZNF300	NM_0528
TLE3	NM_005078	ZNF302	NM_0184
TLN1	NM_006289	ZNF304	NM_0206
TM9SF1	NM_006405	ZNF317	NM_0205
TM9SF2	NM_004800	ZNF338	NM_0220
TMOD2	NM_014548	ZNF345	NM_0034
TMP21	NM_006827	ZNF361	NM_0185
TMSB10	NM_021103	ZNF-U69274	NM_0144
TNFAIP1	NM_021137	ZNRD1	NM_0146
TOMM70A	NM_014820		
TOR2A	NM_130459		
TPT	NM_014317		
TRA1	NM_003299		
TRAF5	NM_004619		
TRAP150	NM_005119		
TRFP	NM_004275		
TRIM4	NM_033017		
TRIP	NM_005879		
TRIP11	NM_004239		
TRN-SR	NM_012470		
TRPS1	NM_014112		
TSG101	NM_006292		
TSLRP	NM_012472		
TSN	NM_004622		
TSNAX	NM_005999		
TUBB4	NM_006086		

Bound Promoters

Reg	Reg	HNF6	HNFAc ₁	HNFAc ₂	HNFAc ₃	HNFAc ₄	HNFAc ₅	HNFAc ₆	HNFAc ₇	HNFAc ₈	HNFAc ₉	HNFAc ₁₀	HNFAc ₁₁	HNFAc ₁₂	HNFAc ₁₃	HNFAc ₁₄	HNFAc ₁₅	HNFAc ₁₆	HNFAc ₁₇	HNFAc ₁₈	HNFAc ₁₉	HNFAc ₂₀	HNFAc ₂₁	HNFAc ₂₂	HNFAc ₂₃	HNFAc ₂₄	HNFAc ₂₅	HNFAc ₂₆	HNFAc ₂₇	HNFAc ₂₈	HNFAc ₂₉	HNFAc ₃₀	HNFAc ₃₁	HNFAc ₃₂	HNFAc ₃₃	HNFAc ₃₄	HNFAc ₃₅	HNFAc ₃₆	HNFAc ₃₇	HNFAc ₃₈	HNFAc ₃₉	HNFAc ₄₀	HNFAc ₄₁	HNFAc ₄₂	HNFAc ₄₃	HNFAc ₄₄	HNFAc ₄₅	HNFAc ₄₆	HNFAc ₄₇	HNFAc ₄₈	HNFAc ₄₉	HNFAc ₅₀	HNFAc ₅₁	HNFAc ₅₂	HNFAc ₅₃	HNFAc ₅₄	HNFAc ₅₅	HNFAc ₅₆	HNFAc ₅₇	HNFAc ₅₈	HNFAc ₅₉	HNFAc ₆₀	HNFAc ₆₁	HNFAc ₆₂	HNFAc ₆₃	HNFAc ₆₄	HNFAc ₆₅	HNFAc ₆₆	HNFAc ₆₇	HNFAc ₆₈	HNFAc ₆₉	HNFAc ₇₀	HNFAc ₇₁	HNFAc ₇₂	HNFAc ₇₃	HNFAc ₇₄	HNFAc ₇₅	HNFAc ₇₆	HNFAc ₇₇	HNFAc ₇₈	HNFAc ₇₉	HNFAc ₈₀	HNFAc ₈₁	HNFAc ₈₂	HNFAc ₈₃	HNFAc ₈₄	HNFAc ₈₅	HNFAc ₈₆	HNFAc ₈₇	HNFAc ₈₈	HNFAc ₈₉	HNFAc ₉₀	HNFAc ₉₁	HNFAc ₉₂	HNFAc ₉₃	HNFAc ₉₄	HNFAc ₉₅	HNFAc ₉₆	HNFAc ₉₇	HNFAc ₉₈	HNFAc ₉₉	HNFAc ₁₀₀	HNFAc ₁₀₁	HNFAc ₁₀₂	HNFAc ₁₀₃	HNFAc ₁₀₄	HNFAc ₁₀₅	HNFAc ₁₀₆	HNFAc ₁₀₇	HNFAc ₁₀₈	HNFAc ₁₀₉	HNFAc ₁₁₀	HNFAc ₁₁₁	HNFAc ₁₁₂	HNFAc ₁₁₃	HNFAc ₁₁₄	HNFAc ₁₁₅	HNFAc ₁₁₆	HNFAc ₁₁₇	HNFAc ₁₁₈	HNFAc ₁₁₉	HNFAc ₁₂₀	HNFAc ₁₂₁	HNFAc ₁₂₂	HNFAc ₁₂₃	HNFAc ₁₂₄	HNFAc ₁₂₅	HNFAc ₁₂₆	HNFAc ₁₂₇	HNFAc ₁₂₈	HNFAc ₁₂₉	HNFAc ₁₃₀	HNFAc ₁₃₁	HNFAc ₁₃₂	HNFAc ₁₃₃	HNFAc ₁₃₄	HNFAc ₁₃₅	HNFAc ₁₃₆	HNFAc ₁₃₇	HNFAc ₁₃₈	HNFAc ₁₃₉	HNFAc ₁₄₀	HNFAc ₁₄₁	HNFAc ₁₄₂	HNFAc ₁₄₃	HNFAc ₁₄₄	HNFAc ₁₄₅	HNFAc ₁₄₆	HNFAc ₁₄₇	HNFAc ₁₄₈	HNFAc ₁₄₉	HNFAc ₁₅₀	HNFAc ₁₅₁	HNFAc ₁₅₂	HNFAc ₁₅₃	HNFAc ₁₅₄	HNFAc ₁₅₅	HNFAc ₁₅₆	HNFAc ₁₅₇	HNFAc ₁₅₈	HNFAc ₁₅₉	HNFAc ₁₆₀	HNFAc ₁₆₁	HNFAc ₁₆₂	HNFAc ₁₆₃	HNFAc ₁₆₄	HNFAc ₁₆₅	HNFAc ₁₆₆	HNFAc ₁₆₇	HNFAc ₁₆₈	HNFAc ₁₆₉	HNFAc ₁₇₀	HNFAc ₁₇₁	HNFAc ₁₇₂	HNFAc ₁₇₃	HNFAc ₁₇₄	HNFAc ₁₇₅	HNFAc ₁₇₆	HNFAc ₁₇₇	HNFAc ₁₇₈	HNFAc ₁₇₉	HNFAc ₁₈₀	HNFAc ₁₈₁	HNFAc ₁₈₂	HNFAc ₁₈₃	HNFAc ₁₈₄	HNFAc ₁₈₅	HNFAc ₁₈₆	HNFAc ₁₈₇	HNFAc ₁₈₈	HNFAc ₁₈₉	HNFAc ₁₉₀	HNFAc ₁₉₁	HNFAc ₁₉₂	HNFAc ₁₉₃	HNFAc ₁₉₄	HNFAc ₁₉₅	HNFAc ₁₉₆	HNFAc ₁₉₇	HNFAc ₁₉₈	HNFAc ₁₉₉	HNFAc ₂₀₀	HNFAc ₂₀₁	HNFAc ₂₀₂	HNFAc ₂₀₃	HNFAc ₂₀₄	HNFAc ₂₀₅	HNFAc ₂₀₆	HNFAc ₂₀₇	HNFAc ₂₀₈	HNFAc ₂₀₉	HNFAc ₂₁₀	HNFAc ₂₁₁	HNFAc ₂₁₂	HNFAc ₂₁₃	HNFAc ₂₁₄	HNFAc ₂₁₅	HNFAc ₂₁₆	HNFAc ₂₁₇	HNFAc ₂₁₈	HNFAc ₂₁₉	HNFAc ₂₂₀	HNFAc ₂₂₁	HNFAc ₂₂₂	HNFAc ₂₂₃	HNFAc ₂₂₄	HNFAc ₂₂₅	HNFAc ₂₂₆	HNFAc ₂₂₇	HNFAc ₂₂₈	HNFAc ₂₂₉	HNFAc ₂₃₀	HNFAc ₂₃₁	HNFAc ₂₃₂	HNFAc ₂₃₃	HNFAc ₂₃₄	HNFAc ₂₃₅	HNFAc ₂₃₆	HNFAc ₂₃₇	HNFAc ₂₃₈	HNFAc ₂₃₉	HNFAc ₂₄₀	HNFAc ₂₄₁	HNFAc ₂₄₂	HNFAc ₂₄₃	HNFAc ₂₄₄	HNFAc ₂₄₅	HNFAc ₂₄₆	HNFAc ₂₄₇	HNFAc ₂₄₈	HNFAc ₂₄₉	HNFAc ₂₅₀	HNFAc ₂₅₁	HNFAc ₂₅₂	HNFAc ₂₅₃	HNFAc ₂₅₄	HNFAc ₂₅₅	HNFAc ₂₅₆	HNFAc ₂₅₇	HNFAc ₂₅₈	HNFAc ₂₅₉	HNFAc ₂₆₀	HNFAc _{261</}
-----	-----	------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	---------------------------

Fig. 20B

Feedforward Loop

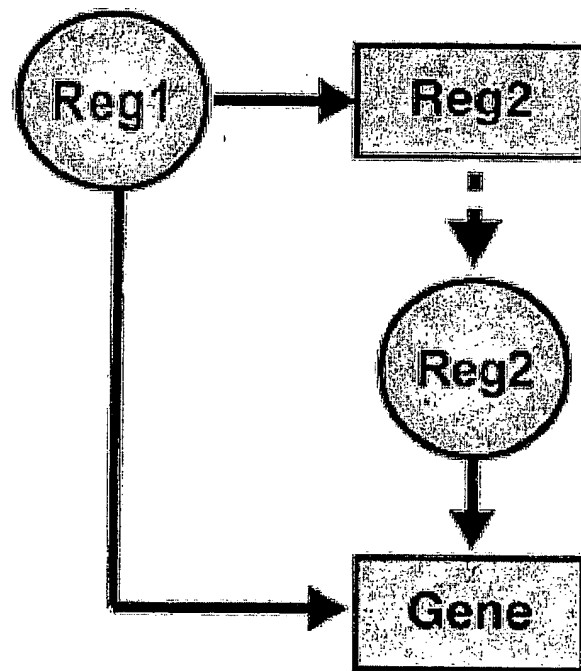


Fig. 21A

Reg1	HNF6		HNF6	
Reg2	HNF4 α		HNF1 α	
Reg3	HNF1 α		HNF1 α	
Bound Promoters	C1S	NM_001734	F11	NM_019559
	ABCC2	NM_000392	C1S	NM_001734
	TNFRSF6	NM_000043	FLJ10650	NM_018168
	UGT2B11	NM_001073	ABCC2	NM_000392
	C2	NM_000063	TNFRSF6	NM_000043
	AMBP	NM_001633	UGT2B11	NM_001073
	SERPING1	NM_000062	UGT1A1	NM_000463
	ADH1B	NM_000668	C2	NM_000063
	PCK1	NM_002591	ADH1A	NM_000667
	DKFZP586A0522	NM_014033	AMBP	NM_001633
	VTN	NM_000638	SERPING1	NM_000062
	AKR1C4	NM_001818	ADH1B	NM_000668
	FLJ21934	NM_024743	HABP2	NM_004132
	KIAA0872	NM_014940	PCK1	NM_002591
	RPL37AP1	NG_000988	DKFZP586A	NM_014033
	PLGL	NM_002665	VTN	NM_000638
	C8B	NM_000066	AKR1C4	NM_001818
	LOC51060	NM_015913	FLJ21934	NM_024743
	HNF4a7	AF509467	KIAA0872	NM_014940
	TM4SF4	NM_004617	RPL37AP1	NG_000988
	UGT2B15	NM_001076	PLGL	NM_002665
	CYP3A43	NM_022820	C8B	NM_000066
	M17S2	NM_031858	LOC51060	NM_015913
	HNMT	NM_006895	HNF4a7	AF509467
	APCS	NM_001639	TM4SF4	NM_004617
	WDR12	NM_018256	UGT2B15	NM_001076
	APOH	NM_000042	CYP3A43	NM_022820
	GJB1	NM_000166	M17S2	NM_031858
	CRP	NM_000567	HNMT	NM_006895
			APCS	NM_001639
			WDR12	NM_018256
			APOH	NM_000042
			GJB1	NM_000166
			CRP	NM_000567

Fig. 21B

Multi-input

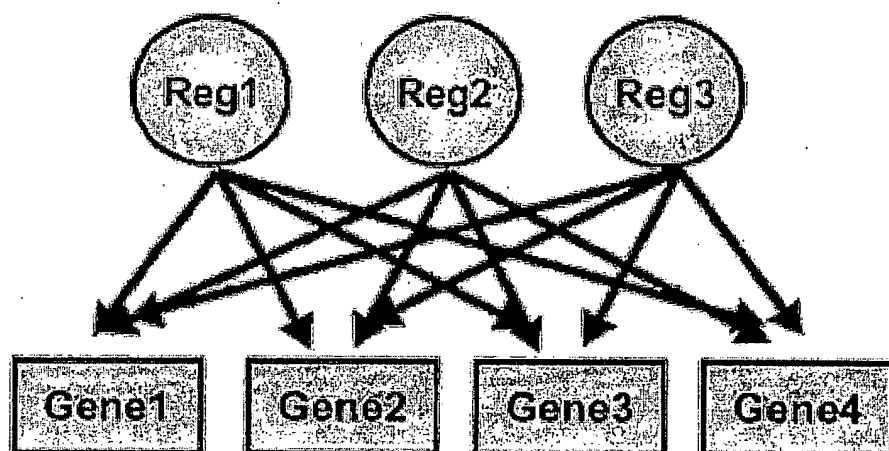


Fig. 22A

Reg1 Reg2	HNF6 HNF4α				HNF1α / HNF4α HNF4α / HNF1α	
Bound Promoters	BCKDHA	NM_000709	FLJ13798	NM_024773	FLJ13273	NM_024751
	FLJ23263	NM_025115	GSS	NM_000178	MGC10500	NM_031477
	FLJ11271	NM_018373	HBOA	NM_007067	SDCCAG10	NM_005869
	HMG2	NM_002129	LOC51060	NM_015913	FBX08	NM_012180
	LOC81558	NM_030802	FLJ13220	NM_021927	ZNF300	NM_052860
	SAS10	NM_020368	FLJ12910	NM_024573	H4F2	NM_003548
	SEC10L1	NM_006544	FLJ10407	NM_018087	FLJ11301	NM_018385
	RRP46	NM_020158	FLJ10342	NM_018064	SEL1L	NM_005065
	SNRPD2	NM_004597	FLJ20671	NM_017924	ZNF155	NM_003445
	MDH1	NM_005917	LOC51287	NM_016565	C6orf11	NM_005452
	ORC1L	NM_004153	GLA	NM_000169	ARHGAP11A	NM_014783
	FLJ20627	NM_017909	RPS6KA5	NM_004755	UROD	NM_000374
	GTF2E1	NM_005513	FLJ20772	NM_017956	FLJ20731	NM_017946
	TOMM70A	NM_014820	FLJ12770	NM_032174	RAB6KIFL	NM_005733
	PAPA-1	NM_031288	FLJ22169	NM_024085	TMP21	NM_006827
	HASJ4442	NM_017528	FLJ10415	NM_018089	MGC15677	NM_032878
	FLJ20084	NM_017659	ZNF317	NM_020933	WBP4	NM_007187
	PEX6	NM_000287	SNW1	NM_012245	PAFAH2	NM_000437
	FLJ11301	NM_018385	REA	NM_007273	EIF3S6	NM_001568
	EED	NM_003797	C2F	NM_006331	PSMA5	NM_002790
	MGC19595	NM_033415	NOLC1	NM_004741	TMOD2	NM_014548
	CIR	NM_004882	CLONE24922	NM_015679	GLA	NM_000169
	CLLD8	NM_031915	CCT8	NM_006585	GNB2L1	NM_006098
	ABCB8	NM_007188	PSMB1	NM_002793	FNTB	NM_002028
	SPG4	NM_014946	WDR12	NM_018256	PEX13	NM_002618
	GABPA	NM_002040	KIAA0806	NM_014813	FE65L2	NM_006051
	OGFR	NM_007346	DKFZp761J139	NM_032280	UQCRC2	NM_003366
	COPB2	NM_004766	SART3	NM_014706	FLJ14855	NM_033210
	AF15Q14	NM_020380	COX7A2L	NM_004718	HHLA2	NM_007072
	MTERF	NM_006980	FLJ20422	NM_017814	CYB5-M	NM_030579
	LOC51633	NM_018023	COPS7A	NM_016319	CDC45L	NM_003504
	FLJ14486	NM_032792	FLJ20643	NM_017916	pcnp	NM_020357
	FLJ21934	NM_024743	HBP1	NM_012257	FLJ20643	NM_017916
	KIAA0872	NM_014940	PSMA1	NM_002786	FLJ21272	NM_025032
	TEGT	NM_003217	FLJ21272	NM_025032		
	MGC4189	NM_032308	FLJ11029	NM_018304		
	SERPINB8	NM_002640	ARL1	NM_001177		
	MGST3	NM_004528	SERPIN1	NM_005025		
	HSP105B	NM_006644	NUDT2	NM_001161		
	C20orf188	NM_015638				

Table S11. The feed forward regulatory motifs in pancreatic islets. The regulatory modules here were derived as described in Supporting Online Material. Feed forwards only involving HNF1α and HNF4α are also multi-input motifs, as they bind each other's promoters in a multicomponent loop.

Fig. 22B

Feedforward Loop

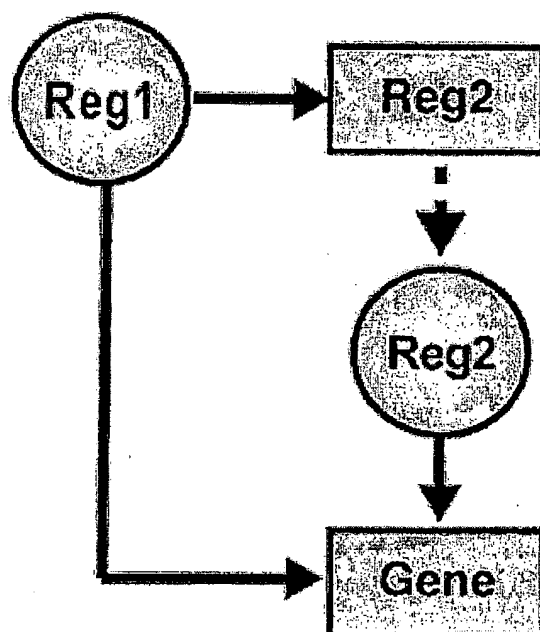


Fig. 23A

Reg1	Reg2	Reg3	HNF1 α HNF6 HNF4 α	HNF1 α HNF6 HNF4 α
Bound Promoters	FLJ10650	NM_018168	FLJ11301	NM_018385
	LOC56906	NM_020147	GLA	NM_000169
	FLJ11301	NM_018385	FLJ20643	NM_017916
	NR0B2	NM_021969	FLJ21272	NM_025032
	KRTAP1.1	NM_030967		
	HNF4a7	AF509467		
	FLJ20156	NM_017691		
	GLA	NM_000169		
	APOH	NM_000042		
	FLJ20643	NM_017916		
	FLJ21272	NM_025032		

Fig. 23B

Multi-input

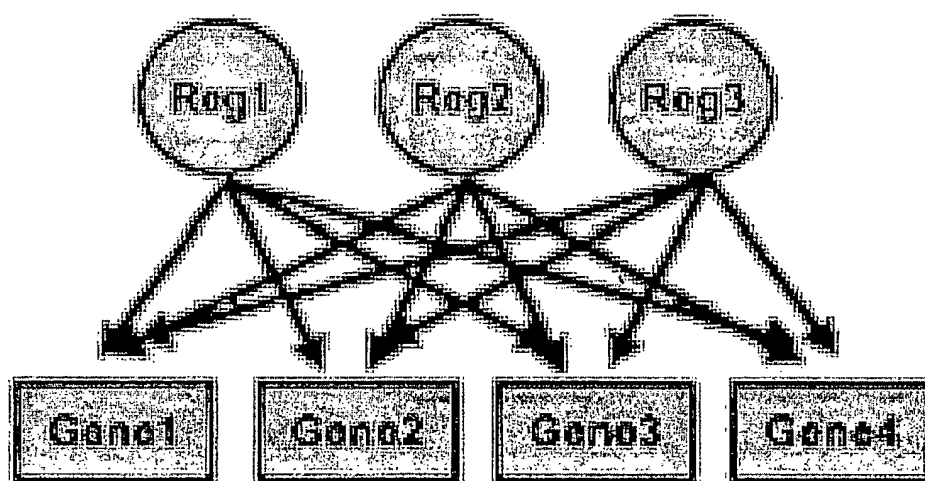


Fig. 24

